

MS FRANCESCA LAMI (Orcid ID : 0000-0002-9706-315X)

Article type : Systematic Review

[Systematic Review: 4 tables; 1 figure; 1 appendix]

Measurement properties of instruments that assess participation in young people with autism spectrum disorder: a systematic review

FRANCESCA LAMI^{1,2}

KRISTINE EGBERTS¹

ALEXANDRA URE^{1,2}

ROWENA CONROY^{2,3,4}

KATRINA WILLIAMS^{1,2,3}

1 Department of Paediatrics, Faculty of MDHS, The University of Melbourne, Melbourne, Victoria; **2** Developmental Disability and Rehabilitation Research, Murdoch Childrens Research Institute, Melbourne, Victoria; **3** Neurodevelopment and Disability, The Royal Children's Hospital, Melbourne, Victoria; **4** School of Psychological Sciences, The University of Melbourne, Melbourne, Victoria, Australia.

Correspondence to Francesca Lami, 50 Flemington Rd Parkville, 3052 Victoria, Australia. Email: francesca.lami@rch.org.au

PUBLICATION DATA

Accepted for publication 00th Month 2017.

Published online 00th Month 2017.

ABBREVIATIONS

ASD Autism spectrum disorder

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/dmcn.13631](https://doi.org/10.1111/dmcn.13631)

This article is protected by copyright. All rights reserved

AYA-ACS	Adolescent and young adult activity card sort
CAPE	Children's activities participation and enjoyment
COSMIN	COnsensus-based Standards for the selection of health Measurement INstruments
ESM	Experience sampling method
ICF	International Classification of Functioning, Disability, and Health
PAC	Preferences for activities of children
PEDI-CAT	Pediatric Evaluation of Disability Inventory, Computer Adaptive Test
PROMIS	Patient-Reported Outcomes Measurement Information System
SFA	School function assessment

AIM To systematically review the measurement properties of instruments assessing participation in young people with autism spectrum disorder (ASD).

METHOD A search was performed in MEDLINE, PsycINFO, and PubMed combining three constructs ('ASD', 'test of participation', 'measurement properties'). Results were restricted to articles including people aged 6 to 29 years. The 2539 identified articles were independently screened by two reviewers. For the included articles, data were extracted using standard forms and their risk of bias was assessed.

RESULTS Nine studies (8 cross-sectional) met the inclusion criteria, providing information on seven different instruments. The total sample included 634 participants, with gender available for 600 (males=494; females=106) and age available for 570, with mean age for these participants 140.58 months (SD=9.11; range=36–624). Included instruments were the school function assessment, vocational index, children's assessment of participation and enjoyment/preferences for activities of children, experience sampling method, Pediatric Evaluation of Disability Inventory Computer Adaptive Test, adolescent and young adult activity card sort, and Patient-Reported Outcomes Measurement Information System parent-proxy peer relationships. Seven studies assessed reliability and validity; good properties were reported for half of the instruments considered. Most studies ($n=6$) had high risk of bias. Overall the quality of the evidence for each tool was limited.

INTERPRETATION Validation of these instruments, or others that comprehensively assess participation, is needed. Future studies should follow recommended methodological standards.

[First page footer]

© Mac Keith Press 2017

DOI: XXXX

[Left page footer]

Developmental Medicine & Child Neurology 2017, 00: 000–000

[Right page footer]

What this paper adds:

- Seven tests have been used to assess participation in young people with autism.
- One test, with excellent measurement properties in one study, does not comprehensively assess participation.
- Studies of three tests that incorporate a more comprehensive assessment of participation have methodological limitations.
- Overall, limited evidence exists regarding measurement properties of participation assessments for young people with autism.

[main text]

Autism spectrum disorder (ASD) is a condition diagnosed by behavioural observation of social communication impairments and the presence of at least two types of restricted, repetitive patterns of behaviour: stereotyped or repetitive motor movements, insistence on sameness or adherence to routines, restricted interests, and/or abnormal reactivity to sensory input. In order to make a diagnosis of ASD, according to Criterion D of the Diagnostic and Statistical Manual of Mental Disorders,¹ symptoms in the social communication and restrictive, repetitive behaviour domains must be associated with impairments in everyday functioning. Examples of such impairments include difficulties with the execution of adaptive daily tasks²⁻⁵ and restrictions in social participation.⁶⁻⁸

The universally recognized model to describe disability is the international classification of functioning, disability, and health (ICF).⁹ Whilst there are many potential benefits associated with using the ICF model and its domains for clinical practice and for translational research, uptake has been slow,^{10,11} and few studies are using this framework with young people with ASD.^{10,12} One suggested reason for the lack of research in this area is that the ICF model lacks conceptual clarity for some of its constructs, such as activity, participation, and environmental factors.^{13,14} However, even if young people with ASD and their parents consider community participation to be one of the most meaningful outcomes to be taken into consideration when assessing the prognosis of individuals with ASD, there is currently no evidence regarding the tools to use to assess participation in this population.¹⁵

Researchers have narrowed the definition of participation to activities that involve the establishment of a role in the community.¹⁴ Whiteneck and Dijkers identify these roles as the establishment or maintenance of a friendship, a family relationship, or a work relationship and the engagement in work, leisure, and spiritual activities, education, employment, community, and civic life.¹⁴ Examples of these roles include being part of structured recreational and/or sport groups, being part of a religious group, attending school or vocational training, having a job, and having a role in human rights associations. Examples of 'activities' that the authors describe include having good conversational skills, changing and maintaining body posture; carrying, moving, and handling objects; performing self-care and house-care chores. Being able to perform acts and tasks that are considered

'activities' may be essential to acquiring independence, but it is the level of societal involvement that is thought to be linked to the attribution of meaning to an individual's life^{14,16,17} and for this reason to their quality of life.^{16,18,19}

Noonan et al. undertook a systematic review of instruments designed for use with adults with ASD,²⁰ and Phillips et al. reviewed the measurement properties of outcome measures of participation in children with a broad range of disabilities.²¹ Both reviews emphasized the importance of measures that consider both quantitative and qualitative aspects of participation. A good measure of participation should be able to assess both (1) extent of involvement, that is, the number of activities that the individual participates in and/or the intensity with which they participate; and (2) satisfaction with/enjoyment of a certain activity by the individual.

Over the past decade, advocates for people with ASD have stressed the importance of focussing on more subjective health-related outcome measures.²²⁻²⁴ In line with the advocates' voices, Morris notes that participation should be understood in relation to a child's capacity, opportunity, and choice.²⁵ In this context, the 'capability approach' describes the person's capacity and opportunities afforded them by the environment, making whether functioning is achieved a matter of personal choice. Yet, in the field of disability, individual choice and preferences are often neglected in studies of participation. Studies that incorporate more subjective measures, such as level of enjoyment of an activity, are needed to understand whether improved participation in social activities equates to a greater quality of life for this population group.

Health-related outcomes should be assessed using instruments that possess adequate measurement properties for the population of interest.^{20,21} In particular, people with ASD are recognized as having unique patterns of strengths and difficulties that could influence the attribution of scores.²⁶⁻²⁸ Their understanding of emotions and capacity to self-reflect may differ from the typically developing population, for example, and this may affect how they rate themselves on particular outcome measures.²⁹ Similarly, interests and enthusiasm may differ in people with ASD and this may influence their attitude towards participation in certain activities. Age and level of intelligence are also critical factors expected to influence an individual's capacity to report on certain experiences³⁰ and their pattern of participation in extracurricular activities.³¹ Studies using well-characterized sample groups that consider the generalisability of their results to the correspondent population of interest are therefore particularly valuable.

Use of parent reported measures to assess participation also requires careful consideration given the subjectivity of personal choice and satisfaction associated with activity and participation. Indeed, discrepancies between child and parent reports have been found when measuring children's preferred activities³¹ and other studies suggest parents underestimate quality of life for their children with disabilities compared with self-report.³² Differences in parent reporting may also be associated with their child's symptom severity. It is possible, for example, that the relative value of participation may be different for parents of nonverbal young adults compared with those who are functioning. To minimize bias, a report by the International Society for Pharmacoeconomics and Outcomes Research³⁰

recommends subjective health experiences should be measured directly from the young person, whenever the individual has the capacity. If this is not possible, observational measures of behaviour are preferred over tools that rely on an inference from a third informant.

Methods to assess the measurement properties of health-related patient reported outcome and the quality of the studies that assess those properties have only recently been agreed upon and are now reported by the COnsensus-based Standards for the selection of Health Measurement INstruments (COSMIN) group. These standards were published in 2010 by the COSMIN group after an online international Delphi study involving clinimetric experts.³³ The COSMIN group also developed a checklist that helps to evaluate the methodological quality of those studies that assess measurement properties for a health-related patient reported outcome instrument.³³

The aim of this systematic review is to answer the question: 'what are the measurement properties of instruments assessing participation in young people with ASD?' We assess whether each identified instrument measures participation, defined as societal involvement,¹⁴ and what domain(s) and dimension(s) of participation each instrument measures. For included instruments we present their measurement properties. Finally we evaluate the methodological quality of each study using the 4-point-rating checklist recommended by the COSMIN group³⁴ and assess overall quality of the evidence.

METHODS

The review protocol has been available on the PROSPERO website, at the following address: www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016040067, since July 2016. However, in August 2017, we modified some of the terms of the search strategy for this systematic review to increase the sensitivity of our search, compared to the original proposed search reported in the published protocol.

Eligibility criteria

Retrieved articles had to meet the following eligibility criteria to be included in this systematic review: (1) reporting on an instrument, such as questionnaires or surveys, interviews, scales, and inventories that assess at least one domain of 'participation'; (2) either being (a) a cohort study, (b) a cross-sectional study investigating the measurement properties of interest in this review (reliability, validity, responsiveness, and interpretability) of the instrument as reported in criterion (1), (c) being a randomized study of instrument accuracy where the participants with ASD were randomized to different instruments, or (d) being a case-control study of instrument accuracy where the participants with ASD and without ASD were evaluated with the same test; (3) investigating these properties in at least one sample of people with ASD aged between 6 and 29 years. We considered only those instruments that aimed to assess participation including some aspects of its narrow definition of societal involvement¹⁴ together with those instruments reported as assessing participation by previous reviews on this topic.^{20,21}

Identification of studies

In August 2017, three databases were searched (MEDLINE, PsycINFO, and Pubmed) to identify studies that looked at the measurement properties of instruments that assess participation, when used with children and young people with ASD (age range 6–29 years). For this search, we followed published recommendations³⁵ and combined the following three constructs and related concepts: (1) our population of interest, which is individuals with ‘autism spectrum disorder’; (2) either the names of instruments identified by others as measuring participation^{20,21} or the combination of terms related to the construct of ‘participation’ as defined in the ICF and the terms of ‘measure’; and finally, (3) the construct of ‘measurement properties’.

After these terms were combined, the articles were filtered for the age range relevant to the review question, namely from school-aged children (minimum age: 6y) to young adults (maximum age: 29y). The upper age range limit was set as equal to the upper age limit for the PsycINFO age group ‘young adulthood’ because this coincides with a critical transition period before adulthood,³⁶ where identity and social role-taking is still well under construction.

For a detailed description of the search strategy, refer to Appendix S1 (online supporting information).

Selection process based on the inclusion criteria

Once duplicates were excluded, two authors independently screened all articles and decided if a study should be included based on the abstract. Each article that passed the initial screen had to meet the inclusion criteria in order to be considered for data extraction. When the reviewers disagreed about the inclusion or exclusion of an article, a third reviewer read the full text and made a final decision on each article.

Data extraction

Data were extracted for each study based on a consistent form. For each article, the first author extracted information regarding: study details, instrument details, and outcomes of interest regarding the measurement properties investigated.

Domains and dimensions of participation

Instrument details include the domains and dimensions of participation investigated by each instrument. Domains of participation are those categories of daily activities that pertain to societal involvement as defined by Whiteneck and Dijkers.¹⁴ Domains were classified according to previously defined categories:^{20,37} ‘leisure activities’; ‘domestic life’; ‘self-care’; ‘interpersonal relationships’; ‘educational activities’; ‘vocational training’; ‘employment’; ‘community life’; ‘religion/spirituality’; ‘economic life’; and ‘caring for others’. Dimensions of participation are those parameters that qualify the daily activities that each person is involved in and usually refer to diversity of activity, intensity of activity, enjoyment of participation, preferences for activities, and with whom and where activities take place.^{31,38} Data was extracted if the instrument measured the ‘extent of involvement in the activity’ and/or the ‘satisfaction/enjoyment after that activity’ because these are considered the main dimensions of participation.^{20,37}

Measurement properties

The measurement properties for data extraction were decided according to the guidelines defined by the COSMIN group.³⁹ The measurement properties included in this review were reliability, validity, and responsiveness. During the data extraction stage, measurement properties were identified in each article and collected as per the guidelines outlined by the COSMIN group.^{39,40} If the study did not use the statistical methods considered standard by the COSMIN group, other information relating to the measurement properties of the instrument was collected. Data pertaining to the generalizability of the results and the interpretability of the included questionnaires were also collected when reported.

Rating of the measurement properties

Two reviewers attributed a qualitative rating as 'positive', 'indeterminate', or 'negative' to the quantitative results extracted for the measurement properties for each tool, using the published criteria for good measurement properties.⁴⁰

Rating of the methodological quality for each study

Each of the nine studies included in this review was assessed by two independent reviewers, for the quality of their methodological design and the risk of bias. The reviewers used the 4-point-rating checklist recommended by the COSMIN group.³⁴ This checklist allows the classification of each study investigating the measurement property of a health-related patient reported outcome measure as poor, fair, good, or excellent quality. This checklist was reported with good interrater agreement and reliability.⁴¹

Overall evidence summary for each tool

The two reviewers created a summary of the quality of the evidence for each tool taking into consideration both the evidence around its measurement properties, the methodological quality of those studies that investigated these measurement properties for that tool, and the content of each tool in terms of domains and dimensions of participation investigated. The reviewers used the recommendations from the COSMIN protocol for psychometric review on measurement properties to provide this overall rating of the evidences.⁴²

RESULTS

Included studies

Figure 1 shows the results of the search and screening process of this systematic review. A total of 2539 studies without duplicates were assessed for eligibility by two independent reviewers. Full text review was completed for 44 articles, independently by two authors. Of these articles, only nine met the inclusion criteria for this study.

The remaining 35 articles were excluded for two reasons. Twenty-five articles were excluded because the instrument used did not assess the construct of participation; rather the instrument evaluated activity, often adaptive behaviour, or global everyday functioning. A further 10 articles were excluded because the participants did not meet the eligibility criteria because of diagnosis (for example, one study included all types of developmental disabilities) or age range (e.g. one study focused on preschool children only).

Eight out of the nine included studies were cross-sectional in design^{26,43–49} and one was a case control study, using a known-group method.⁵⁰ Overall, the age range of the sample investigated was from 3 to 52 years (weighted mean for seven studies with a total of 570 participants=140.58m, SD=9.11m), and three studies only included adolescents (age range 10–25y; mean=16y 9m).^{45–47} Three studies included people with ASD with and without intellectual disability,^{26,45,49} and four included people with ASD only.^{43,44,46,47} Table I reports some information regarding the samples characteristics for the studies included. Seven studies had author teams based in the United States of America^{26,45–50} and all used instruments published in English. The setting where the instrument was administered varied between studies and included school,⁵⁰ online testing,^{26,45,49} and a research environment.^{46–48} Two studies used real-time online surveys via electronic devices and, therefore, the individuals' experiences were measured both in their home environment and in their community.^{43,44} Four studies included tools that were completed by a third person, usually a parent or caregiver,^{26,45,48–50} and two studies used tools that were self-completed by the individual.^{43,44,46} In one study, for a subset of participants (46.67%), both the parent and child completed the checklist in order to test the parent-child's agreement and/or to collect factual information about the activities when the child was not able to answer themselves.⁴⁷ Three studies used computer adaptive technology.^{26,45,49} The sample size ranged from 11⁵⁰ to 365²⁶ participants, and five studies had 30 or less participants.^{43,44,46,47,50}

Instruments included in this systematic review

From the nine included studies, seven instruments were identified, with two different versions of one instrument. The instruments assessed were the school function assessment (SFA), the vocational index for adults with autism, the children's activities participation and enjoyment (CAPE)/preferences for activities of children (PAC), the experience sampling method (ESM), the adolescents and young adult activity card sort (AYA-ACS), the Patient-Reported Outcomes Measurement Information System (PROMIS) parent-proxy peer relationships measure, and the Pediatric Evaluation of Disability Inventory, Computer Adaptive Test (PEDI-CAT) and its ASD version.

Domains and dimensions of participation

The CAPE/PAC, the AYA-ACS, and the ESM were specifically designed to evaluate societal participation. These were also the only instruments that investigated both the extent and the enjoyment of involvement by the individual^{26,44,46,47} and were self-reported.

The PROMIS parent-proxy peer relationships measure and the vocational index were included in this systematic review because they measure interpersonal relationships (the PROMIS) and vocational activities or employment (vocational index), which are domains of participation.¹³

The SFA and the PEDI-CAT are designed to provide a global measure of daily functioning, in the school context (SFA) and broadly in everyday life (PEDI-CAT). Each contains a subscale assessing one dimension of participation. That is, the SFA assesses the extent of participation in six school activities. In the PEDI-CAT there is a subscale named 'Responsibility' that assesses the overall responsibility, also thought of as extent of involvement, in managing a certain task in that life situation. Details about the tools that were assessed in each study are described in Table I.

Measurement properties

Seven out of nine studies^{26,43–47,49} investigated the reliability (either test-retest reliability, interrater reliability, internal consistency, or standard error of measurement) of the instrument of interest, together with either its construct or its content validity. Table II reports a summary of the measurement property data extracted for each study. No studies investigated responsiveness, according to the COSMIN standards. The measurement properties (reliability and validity) of each tool were rated as ‘positive/+’, ‘indeterminate/?’, or ‘negative/-’ according to the consensus of two independent reviewers using the criteria established by Terwee et al.⁴⁰ (see Table II).

Although some of the studies investigated the same broad measurement properties, they often looked at different aspects of that property. For example, some studies looked at the test-retest reliability of the tool, but they did not look at the tool interrater reliability^{43–45} or vice versa.⁴⁶ In addition, these studies often used different statistical techniques to test the measurement properties. Therefore, the reviewers were not able to make quantitative comparisons between tools with respect to reliability and validity.

Instrument performance

The CAPE and the PAC measurement properties were assessed separately. The CAPE demonstrated good test-retest reliability for almost all the subscales ($r > 0.70$) when used with adolescents with high functioning ASD. However, the test-retest reliability of the PAC ($r < 0.70$) was not found to be sufficient according to the COSMIN standards. Potvin et al.⁴⁷ evaluated content validity of the CAPE/PAC by determining if the children reported participating in recreational activities other than the 55 included in this tool. The authors declared that this tool covered a broad range of recreational activities in which children with high functioning ASD participate in; however they also recommend routinely asking the young person if they participated in additional activities that were not included in the questionnaire. Four of 14 children named additional activities, with two (‘inventing things’ and ‘going online’) activities that had not been included in the tool subscales. In addition, 12 of 14 children who participated in the content validity study of the PAC named 22 other activities that they would prefer to do which were not listed. These included activities related to the child’s specific areas of interests (e.g. ‘actual Pokemon training’). There was no assessment of the content validity compared to the control sample without ASD. Finally, to assess parent and child agreement, parents ($n=30$) were asked if they agreed with their child’s rating. Twelve of 16 (75%) and 7 of 14 (50%) parents agreed with their children on the CAPE and the PAC respectively. However, this method is not considered a standard to assess interrater reliability by the COSMIN group.

The test-retest reliability and validity of the ESM were not provided in accordance with COSMIN standards in the two studies that assessed this tool and therefore it is not possible to provide a qualitative rating of the measurement properties of the ESM.

In addition, the study that looked at the interrater reliability between young people’s and their parents’ scores on the AYA-ACS did not perform analysis using the COSMIN standards, therefore the

reviewers could not determine the rating after the recommendations provided by Terwee et al.⁴⁰ However, evidence pertaining to the validity of the AYA-ACS received a positive rating (see Table II).

The PROMIS parent proxy peer relationships measure was rated as showing good internal consistency (see Table II). The content validity of the PROMIS parent proxy peer relationships measure was measured by asking five parents to participate in a semi-structured interview about the relevance of this scale for their children. The study reported that parents considered the questions from the scale to be relevant for their children, but noted that their responses to these items would vary from week to week. Furthermore, the reviewers rated the evidences regarding concurrent validity as positive (there was an expected inverse relationship between autism severity and overall score in the PROMIS measure of peer relationship, $r=-0.60$, $p<0.01$).

The study on the vocational index did not aim to analyse the psychometric properties, but the authors reported that when they performed an analysis of the category agreement by two independent reviewers for 10 per cent of the participants, this agreement was excellent at 94 per cent (Cronbach's $\alpha=0.92$).

The responsibility scale of the PEDI-CAT, both for the original and the ASD version, had excellent validity (see Table II). Although the responsibility scale of the PEDI-CAT demonstrated good test-retest reliability ($r=0.90$), the study⁴⁵ about this measurement property was done with a small sample. For the ASD version of the same tool, Coster et al.²⁶ found excellent internal consistency (acceptable fit to the model).

There was no assessment of reliability for the SFA, but the results regarding the know group analysis by pairwise comparison showed that the SFA participation subscale scores were able to differentiate those children with a diagnosis of ASD to those with a diagnosis of learning disability ($X^2=11.98$, $p=0.002$).

Rating of the methodological quality for each study

Table III presents a summary of the rating of the methodological quality for the studies included in this review, specifying for each study the main risk of bias.

The study about the PEDI-CAT, ASD version had low risk of bias.²⁶ Two-thirds of the studies employed a small size for their type of assessment,^{43–47,50} and were rated as having a high risk of bias and a poor methodological quality. As protocols were not available, the reporting bias is unclear for five studies.^{26,43,44,46,50} However, all of the included studies reported outcomes that are appropriate to the measurement assessment.

Quality of the evidence for each tool

Table IV presents a summary of the evidence regarding each of the identified tools. For seven tools included in this review this quality summary is based on one study. Table IV also includes a summary of the domains and dimensions of participation investigated by each tool.

This systematic review identified moderate evidence for the measurement properties of the PEDI-CAT, ASD version when used with individuals with ASD aged between 3 and 21 years old. In

addition, there is limited evidence about the validity of the PEDI-CAT for its use with adolescents with ASD. However, there are limitations with respect to the scope of this measure, in that it focuses on the extent of involvement in some activities of daily life, without capturing more subjective aspects of participation. There is limited or insufficient evidence available to comprehensively assess the measurement properties of the majority of the tools included in this systematic review. In particular, there was a lack of evidence available to assess the measurement properties of the CAPE/PAC and AYA-ACS which were the only tools to incorporate multiple domains and dimensions of participation.

DISCUSSION

Principal findings

Participation in society is considered a valuable outcome measure in reference to young people with ASD and their parents. Recently there has been increasing recognition of the importance of measuring both involvement in an activity and personal satisfaction associated with an individual's level of participation, in order to understand relationships between social participation and improvement in quality of life.^{14,20,21} A past systematic review of tools used in ASD early intervention and observational studies used a broader definition of outcomes in younger children.¹⁵ McConachie et al.¹⁵ noted that participation in life situations was highly valued by parents of children with ASD but were unable to identify a robust tool for assessing this specific outcome. Focussing on an older sample of individuals with ASD (aged 6–29y) and a more specific search strategy for participation, our systematic review identified nine studies that have assessed the measurement properties of instruments evaluating participation. However, close inspection of the content of these instruments revealed variability in how participation is measured, with only a subset of studies incorporating a young person's personal choice and level of satisfaction in their level of participation. Although most studies reported adequate measurement properties for the instrument investigated, the majority also had a high risk of bias and therefore the quality of evidence for the tool was limited.^{26,43–50}

The responsibility scale of the PEDI-CAT, ASD version, demonstrated excellent reliability and validity when used with young people with ASD, suggesting that it is an effective tool for measuring parents' reporting on the extent to which the young person is managing life tasks. However, given the tool does not measure the individual's level of enjoyment or satisfaction in the task, and rather measures the degree of difficulty associated with performing the task, it could be argued that the subscale is not a comprehensive measure of participation as defined by the ICF and other experts in the field.^{20,21}

The CAPE/PAC and the AYA-ACS are self-report tools and their use with older individuals with ASD provides the opportunity to collect information on subjective dimensions of enjoyment associated with participation. In addition, the preferences for activities of children (PAC) offers an opportunity to document these individuals' preferences for the activities they would like to be involved in, which allow clinicians and researchers to understand individuals' meaningful level of participation without simply superimposing a model of optimal or 'normal' participation to every individual.²⁵ Not surprisingly, only half of the parents interviewed by Potvin et al.⁴⁷ agreed with the rating provided by their children in the PAC. This is crucial because it highlights the importance of investigating the

individual with ASD's point of view when assessing preferences for activities. Parent and child report on other tools also vary, usually because a parent does not know what their child is thinking. With regards to participation in autism, some people with ASD may not enjoy certain specific activities in the same way of people who do not have an ASD diagnosis⁵¹ and this may influence whether they want to participate or not in certain activities. Furthermore, in addition to the participation tool, investigating areas of special interests using open questions would be recommended in this population.⁴⁷ Overall, despite the essential information about participation provided by the CAPE/PAC and AYA-ACS, there is limited evidence about the measurement properties of these tools when used with young individuals with ASD.

The SFA assesses the young person's involvement in school activities but does not include a subjective dimension of participation. Attending school is linked to developing independence and providing opportunities to understand cultural and societal norms.⁵² For these reasons, the SFA could be a valuable instrument; however further research investigating both the reliability and validity of the test using a larger sample size is needed to provide more confidence in its use with young people with ASD.

The vocational index was developed to create a comprehensive index of vocational education and employment for adults with ASD, given the importance that this domain of societal participation has in relation to the development of independence and wealth during adulthood. The main aim of Taylor and Seltzer's study⁴⁸ was to describe the development of this tool, but this study also offered some information regarding the vocational index interrater reliability and less systematic information on its validity. In addition, Taylor and Seltzer⁴⁶ state that this test should be used 'in conjunction with other measures of successful adulthood' and should be cross-culturally validated, as some of the categories may be quite specific for the socio-geographical areas where this tool was developed.

Young people with ASD can experience various levels of inclusion in peers circles and take on more or less active roles in those relationships.⁵³ However, the research in this area is still quite limited and in order to advance in this field, a reliable and valid assessment to capture differences in involvement in social relationships would be beneficial. The PROMIS parent-proxy peers relationships measure would be able to fill this gap, but as mentioned by Toomey et al.⁴⁹ further research to assess its content validity and responsiveness is needed.

Finally, the ESM is considered to be an ecological measure of participation. The use of an experience sampling method such as the ESM minimizes recall bias and allows for a study of the links between participation and elicited emotions. However, the evidences regarding the measurement properties of the ESM when used with individuals with ASD are insufficient.^{43,44}

Strengths and weaknesses

This systematic review adds to previous research that identified a lack of evaluated measures of participation in young children with ASD.¹⁵ Our focus was on tools used to assess participation with young people with ASD aged between 6 and 29 years old. In contrast to McConachie et al.¹⁵ the current study found seven tools that measure participation and had their measurement properties assessed in samples of young people with ASD. The differences in the findings may be because of our focus on an older age range where possibly there is a major interest on functioning outcome

measures, compared to diagnostic outcome measures. It is also possible that our systematic review used a more sensitive search strategy in relation to the terms related to participation. In line with the recommendations by McConachie et al.¹⁵ we used the study-design-relevant COSMIN methods to assess the methodology and risk of bias for each study and quality of the evidence for each tool.

Whilst decisions about which articles to include in this review were based on defined guidelines and definitions from previous literature,^{14,20,21} it is likely that a degree of subjectivity associated with the construct of participation influenced the decisions of the two independent reviewers. Decisions were complicated by the multiple domains and dimensions of participation which were not always clearly defined in previous studies.^{26,45,47,49,50} Greater clarity in the definition of participation would improve future studies assessing participation instrument performance.²¹

Clinical and research implications and future directions

Participation is now widely accepted as important for individuals, carers, clinicians, service providers, and policy makers. Despite it being broadly recognized that societal participation should be assessed as a meaningful outcome measure of individuals with ASD, it is still undefined what tool should be used with individuals with ASD.

Existing tools that were evaluated in this population fall short of what is needed for a comprehensive assessment of participation. Evidence about their measurement properties is also incomplete and most studies had a high risk of bias.

To assess participation in a way that could benefit young people with ASD we need instruments that fulfill requirements described by experts in the field^{20,21,37} and remain relevant to the ICF framework. Clinicians and researchers needing to choose a tool for this purpose should be cautious when using one of the tools identified in this systematic review.

Unanswered questions and implications for future research and clinical practice

Maximizing effective participation in society is considered an overarching goal of services for young people with ASD. As meaningful, comprehensive assessments of participation incorporate personal preferences and individual enjoyment, self-reporting tools are deemed most appropriate. However, to date, no tool is considered an optimal measure of participation for individuals with ASD. Results of this systematic review suggests further examinations of the measurement properties of existing tools or the development of new tools specific for this population is required.

Researchers developing and evaluating patient-reported assessment measures should consider the variability that exists within the autism spectrum, and consider their participants' ability to report feelings and experiences and functional level of their social communication. The different ways people with ASD interpret and value certain items should also be considered as identified by other authors in relation to quality of life²⁷ and anxiety.²⁸ In particular, we recommend that when measuring the reliability or content validity of a subjective construct, such as enjoyment and preferences, these properties should be investigated from the individuals' perspectives, rather than the perspective of their caregivers.

We also recommend validation and/or development of patient-reported tools using pictorial information to facilitate the expression of those with more severe social communication impairments

and concrete thinking, to advance opportunities for understanding important outcomes for all individuals with ASD. However, if researchers and clinicians would like to investigate more objective and/or specific domains of participation, such as vocational activities, peer relationships, or curricular or extracurricular participation, the involvement of an informant would be considered appropriate if reporting on a direct observation of behaviour in accordance with the recommendations from the International Society for Pharmacoeconomics and Outcomes Research task force report.³⁰

Authors who wish to validate tools for their use in populations of individuals with ASD, should consider using the freely available resources from the COSMIN group (www.cosmin.nl) or the resources provided by the Patient Reported Outcomes Measurement Group at the University of Oxford (<http://phi.uhce.ox.ac.uk/home.php>), as suggested by McConachie et al.¹⁵

ACKNOWLEDGMENTS

This review was completed during the PhD of Ms Francesca Lami. Her PhD is funded through an Australian Postgraduate Award and through a top-up studentship from the Lorenzo and Pamela Galli Foundation. Drs Kristine Egberts and Alexandra Ure are postdoctoral fellows within the Department of Paediatrics, whose positions are funded by the Lorenzo and Pamela Galli Foundation. Professor Katrina Williams and Dr Rowena Conroy are PhD supervisors for Francesca Lami, and work for the organizations listed above. All the authors declare that they have no conflict of interests in publishing the results of this study.

The authors would like to acknowledge and thank Ms Poh Chua for her support during the development of the search strategy and Ms Amanda Brignell for the feedback provided on the initial draft of this article.

SUPPORTING INFORMATION

The following additional material may be found online:

Appendix S1: Example of search terms used in the PubMed search

REFERENCES

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th edn. Washington, DC: Author, 2013.
2. Matthews NL, Smith CJ, Pollard E, Ober-Reynolds S, Kirwan J, Malligo A. Adaptive functioning in autism spectrum disorder during the transition to adulthood. *J Autism Dev Disord* 2015; **45**: 2349–60.
3. Pugliese CE, Anthony L, Strang JF, Dudley K, Wallace GL, Kenworthy L. Increasing adaptive behavior skill deficits from childhood to adolescence in autism spectrum disorder: role of executive function. *J Autism Dev Disord* 2015; **45**: 1579–87.
4. Pugliese CE, Anthony LG, Strang JF, et al. Longitudinal examination of adaptive behavior in autism spectrum disorders: influence of executive function. *J Autism Dev Disord* 2016; **46**: 467–77.
5. Duncan AW, Bishop SL. Understanding the gap between cognitive abilities and daily living skills in adolescents with autism spectrum disorders with average intelligence. *Autism* 2015; **19**: 64–72.

6. Orsmond GI, Krauss MW, Seltzer MM. Peer relationships and social and recreational activities among adolescents and adults with autism. *J Autism Dev Disord* 2004; **34**: 245–56.
7. Orsmond GI, Kuo HY. The daily lives of adolescents with an autism spectrum disorder: discretionary time use and activity partners. *Autism* 2011; **15**: 579–99.
8. Shattuck PT, Orsmond GI, Wagner M, Cooper BP. Participation in social activities among adolescents with an autism spectrum disorder. *PLoS One* 2011; **6**: e27176.
9. World Health Organization. International classification of functioning, disability and health (ICF). Geneva: World Health Organization, 2001.
10. Bruyère SM, Van Looy SA, Peterson DB. The International Classification of Functioning, Disability and Health: contemporary literature overview. *Rehabilitation Psychology* 2005; **50**: 113–21.
11. de Schipper E, Mahdi S, de Vries P, et al. Functioning and disability in autism spectrum disorder: a worldwide survey of experts. *Autism Res* 2016; **9**: 959–69.
12. Ibragimova N, Granlund M, Björck-Åkesson E. Field trial of ICF version for children and youth (ICF-CY) in Sweden: logical coherence, developmental issues and clinical use. *Dev Neurorehabil* 2009; **12**: 3–11.
13. Badley EM. Enhancing the conceptual clarity of the activity and participation components of the International Classification of Functioning, Disability, and Health. *Soc Sci Med* 2008; **66**: 2335–45.
14. Whiteneck G, Dijkers MP. Difficult to measure constructs: conceptual and methodological issues concerning participation and environmental factors. *Arch Phys Med Rehabil* 2009; **90**(Suppl. 11): S22–35.
15. McConachie H, Parr JR, Glod M, et al. Systematic review of tools to measure outcomes for young children with autism spectrum disorder. *Health Technol Assess* 2015; **19**: 1–506.
16. Tobin MC, Drager KD, Richardson LF. A systematic review of social participation for adults with autism spectrum disorders: Support, social functioning, and quality of life. *Res Autism Spectr Disord* 2014; **8**: 214–29.
17. Rosenbaum P, Gorter JW. The ‘F-words’ in childhood disability: I swear this is how we should think! *Child Care Health Dev* 2012; **38**: 457–63.
18. Billstedt E, Gillberg IC, Gillberg C. Aspects of quality of life in adults diagnosed with autism in childhood: A population-based study. *Autism* 2011; **15**: 7–20.
19. García-Villamizar D, Wehman P, Navarro MD. Changes in the quality of autistic people's life that work in supported and sheltered employment. A 5-year follow-up study. *J Vocat Rehabil* 2002; **17**: 309–12.
20. Noonan VK, Kopec JA, Noreau L, Singer J, Dvorak MF. A review of participation instruments based on the International Classification of Functioning, Disability and Health. *Disabil Rehabil* 2009; **31**: 1883–901.
21. Phillips RL, Olds T, Boshoff K, Lane AE. Measuring activity and participation in children and adolescents with disabilities: a literature review of available instruments. *Aust Occup Ther J* 2013; **60**: 288–300.
22. Kapp SK, Gillespie-Lynch K, Sherman LE, Hutman T. Deficit, difference, or both? Autism and neurodiversity. *Dev Psychol* 2013; **49**: 59–71.

23. Pellicano E, Stears M. Bridging autism, science and society: moving toward an ethically informed approach to autism research. *Autism Res* 2011; **4**: 271–82.
24. Robertson SM. Neurodiversity, quality of life, and autistic adults: shifting research and professional focuses onto real-life challenges. *Disabil Stud Q* 2010; **30**. Available at: <http://dsq-sds.org/article/view/1069>
25. Morris C. Measuring participation in childhood disability: how does the capability approach improve our understanding? *Dev Med Child Neurol* 2009; **51**: 92–4.
26. Coster WJ, Kramer JM, Tian F, et al. Evaluating the appropriateness of a new computer-administered measure of adaptive function for children and youth with autism spectrum disorders. *Autism* 2016; **20**: 14–25.
27. Tavernor L, Barron E, Rodgers J, McConachie H. Finding out what matters: validity of quality of life measurement in young people with ASD. *Child Care Health Dev* 2013; **39**: 592–601.
28. Wigham S, McConachie H. Systematic review of the properties of tools used to measure outcomes in anxiety intervention studies for children with autism spectrum disorders. *PLoS ONE* 2014; **9**: e85268.
29. Attwood T. The complete guide to Asperger's syndrome: London; Philadelphia: Jessica Kingsley Publishers, 2007.
30. Matza LS, Patrick DL, Riley AW, et al. Pediatric patient-reported outcome instruments for research to support medical product labeling: report of the ISPOR PRO good research practices for the assessment of children and adolescents task force. *Value Health* 2013; **16**: 461–79.
31. King G, Law M, Hurley P, Petrenchik T, Schwellnus H. A developmental comparison of the out-of-school recreation and leisure activity participation of boys and girls with and without physical disabilities. *Int J Disabil Dev Educ* 2010; **57**: 77–107.
32. White-Koning M, Grandjean H, Colver A, Arnaud C. Parent and professional reports of the quality of life of children with cerebral palsy and associated intellectual impairment. *Dev Med Child Neurol* 2008; **50**: 618–24.
33. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010; **19**: 539–49.
34. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012; **21**: 651–7.
35. Terwee CB, Jansma EP, Riphagen II, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009; **18**: 1115–23.
36. Arnett JJ. Emerging adulthood: What is it, and what is it good for? *Child Dev Perspect* 2007; **1**: 68–73.
37. McConachie H, Colver AF, Forsyth RJ, Jarvis SN, Parkinson KN. Participation of disabled children: how should it be characterised and measured? *Disabil Rehabil* 2006; **28**: 1157–64.

- 38.** Bellamy G, Newton J, LeBaron N, Horner R. Quality of life and lifestyle outcomes: a challenge for residential programs. In: Schalock RL, editor. *Quality of life: Perspectives and issues*. Washington DC, USA: American Association on Mental Retardation, 1990: 127–37.
- 39.** Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; **63**: 737–45.
- 40.** Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; **60**: 34–42.
- 41.** Mokkink LB, Terwee CB, Gibbons E, et al. Inter-rater agreement and reliability of the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) checklist. *BMC Med Res Methodol* 2010; **10**: 82.
- 42.** Terwee CB dVH, Prinsen CAC, Mokkink LB. Protocol for systematic reviews of measurement properties [Internet]. Amsterdam: VU University Medical Center, 2011: 15. Available at: [http://www.cosmin.nl/images/upload/files/Protocol%20klinimetriche%20review%20version%20nov%202011\(1\).pdf](http://www.cosmin.nl/images/upload/files/Protocol%20klinimetriche%20review%20version%20nov%202011(1).pdf) (accessed 10 July 2017).
- 43.** Chen YW, Bundy A, Cordier R, Einfeld S. Feasibility and usability of experience sampling methodology for capturing everyday experiences of individuals with autism spectrum disorders. *Disabil Health J* 2014; **7**: 361–6.
- 44.** Chen YW, Cordier R, Brown N. A preliminary study on the reliability and validity of using experience sampling method in children with autism spectrum disorders. *Dev Neurorehabil* 2015; **18**: 383–9.
- 45.** Kramer JM, Liljenquist K, Coster WJ. Validity, reliability, and usability of the Pediatric Evaluation of Disability Inventory-Computer Adaptive Test for autism spectrum disorders. *Dev Med Child Neurol* 2016; **58**: 255–61.
- 46.** McCollum M, LaVesser P, Berg C. Participation in daily activities of young adults with high functioning autism spectrum disorder. *J Autism Dev Disord* 2016; **46**: 987–97.
- 47.** Potvin MC, Snider L, Prelock P, Kehayia E, Wood-Dauphinee S. Children's assessment of participation and enjoyment/preference for activities of children: psychometric properties in a population with high-functioning autism. *Am J Occup Ther* 2013; **67**: 209–17.
- 48.** Taylor JL, Seltzer MM. Developing a vocational index for adults with autism spectrum disorders. *J Autism Dev Dis* 2012; **42**: 2669–79.
- 49.** Toomey M, Schwartz J, Laverdiere M, et al. Preliminary validation of the PROMIS Parent-Proxy Peer Relationships Measure in children with autism spectrum disorder: a DBPNet study. *J Dev Behav Pediatr* 2016; **37**: 724–9.
- 50.** Davies PL, Soon PL, Young M, Clausen-Yamaki A. Validity and reliability of the school function assessment in elementary school students with disabilities. *Phys Occup Ther Pediatr* 2004; **24**: 23–43.
- 51.** Eversole M., Collins DM, Karmarkar A, et al. Leisure activity enjoyment of children with autism spectrum disorders. *J Autism Dev Dis* 2016; **46**: 10–20.

52. Simeonsson RJ, Carlson D, Huntington GS, McMillen JS, Brent JL. Students with disabilities: a national survey of participation in school activities. *Disabil Rehabil* 2001; **23**: 49–63.
53. Calder L, Hill V, Pellicano E. ‘Sometimes I want to play by myself’: understanding what friendship means to children with autism in mainstream primary schools. *Autism* 2013; **17**: 296–316.
54. Coster WJ, Deeney TA, Haltiwanger JT, Haley SM. School function assessment. Boston, MA: Boston University, 1998.
55. King G, Law M, King S, et al. Children's assessment of participation and enjoyment (CAPE) and preferences for activities of children (PAC). San Antonio, TX: Harcourt Assessment, Inc., 2004.
56. Hektner JM, Schmidt JA, Csikszentmihalyi M. Experience sampling method: measuring the quality of everyday life. Thousand Oaks, CA; London: SAGE, 2007.
57. Haley SM, Coster WJ, Dumas HM, et al. Pediatric evaluation of disability inventory computer adaptive test. CRE care [Internet]. Available at: <https://www.pedicat.com/> (accessed 10 December 2016).
58. Kramer JM, Coster WJ, Kao YC, Snow A, Orsmond GI. A new approach to the measurement of adaptive behavior: development of the PEDI-CAT for children and youth with autism spectrum disorders. *Phys Occup Ther Pediatr* 2012; **32**: 34–47.
59. Dewalt DA, Thissen D, Stucky BD, et al. PROMIS pediatric peer relationships scale: development of a peer relationships item bank as part of social health measurement. *Health Psychol* 2013; **32**: 1093–103.
60. Berg C, McCollum M, Cho E, Jason D. Development of the adolescent and young adult activity card sort. *OTJR (Thorofare N J)* 2015; **35**: 221–31.
61. Haley SM, Coster WJ, Dumas HM, Fragala-Pinkham MA, Moed R. Pediatric Evaluation of Disability Inventory Computer Adaptive Test (PEDI-CAT). Development, standardization and administration manual. Boston University Medical Center, Boston, MA: Trustees of Boston University, 2015.

[Figure legends]

Figure 1: Flow chart of the study inclusion

Table I: Domains and dimension of participation as investigated by each instrument

Study ^a	Instrument administered	Type of instrument and informer	Domain of participation investigated	Dimension of participation investigated	Sample size	Age range in years	% males	Characteristics of the sample	Country where the study was done
Davies et al. ⁵⁰	SFA ⁵⁴	teacher or therapist completed scale	educational activities	not reported	11	6–11	73%	ns	USA
Taylor and Seltzer ⁴⁸	vocational index ⁴⁸	index provided via interview with parent	vocational activities and employment	extent of involvement	34	18–52	ns	ns	USA
Potvin et al. ⁴⁷	CAPE/PAC ⁵⁵	self-reported (with help of caregiver, if appropriate) checklist	leisure activities; interpersonal relationships; domestic life; educational activities; community life; religion/spirituality	extent of involvement; enjoyment or satisfaction	30	7–13	86.6%	only participants without ID	USA
Chen et al. ^{43,44}	ESM ⁵⁶	self-reported ecological momentary assessment	domestic life (household task), interpersonal relationship, leisure activities,	extent of involvement; enjoyment or satisfaction	4 6	16–32 8–12	50% 100%	only participants without ID	Australia

			educational activities						
Kramer et al. ⁴⁵	PEDI-CAT ⁵⁷	caregiver-reported scale (computerized adaptive testing)	*	extent of involvement	39	10–18	87.2%	participants with and without ID	USA
Coster et al. ²⁶	PEDI-CAT, ASD version ⁵⁸	caregiver-reported scale (computerized adaptive testing)	*	extent of involvement	365	3–21	83.3%	participants with and without ID	USA
McCollum et al. ⁴⁶	AYA-ACS ⁶⁰	self-reported cards sort	domestic life (household task), caring for others, interpersonal relationship, leisure activities, educational activities, vocational training and employment	extent of involvement; enjoyment or satisfaction	24	18–25	71%	only participants without ID	USA

Toomey et al. ⁴⁹	PROMIS parent-proxy peer relationships measure ⁵⁹	caregiver-reported scale (computerized adaptive testing)	interpersonal relationships, community life and caring for others	extent of involvement	121	5–12	86.8%	participants with and without ID	USA
-----------------------------	--	--	---	-----------------------	-----	------	-------	----------------------------------	-----

^aThe only information that the authors of this review could find regarding the domain of participation assessed by the responsibility scale were those listed on the PEDI-CAT website, where it was stated that the domains assessed by this subscale are 'organization and planning, taking care of daily needs, health management and staying safe'.⁵⁷ SFA, school function assessment; ns, not specified; CAPE, children's assessment of participation and enjoyment; PAC, preferences for activities of children; ID, intellectual disability; ESM, experience sampling method; PEDI-CAT, Pediatric Evaluation of Disability Inventory, Computer Adaptive Test; ASD, autism spectrum disorder; AYA-ACS, adolescents and young adults activity card sort; PROMIS, Patient-Reported Outcomes Measurement Information System.

Table II: Measurements properties of instruments

Study	Instrument administered	Reliability (rating positive +/negative -/ indeterminate ?) <i>Data extracted on statistical analysis (statistical value found)</i>	Validity (rating positive +/negative -/ indeterminate ?) <i>Data extracted on statistical analysis (statistical value found)</i>
-------	-------------------------	--	---

Davies et al. ⁵⁰	SFA ⁵	Not assessed for the ASD group	Construct validity (+) <i>Known group validity: Kruskal-Wallis analysis ($X^2=11.98$; ASD vs LD $p=0.002$)</i>
Taylor and Seltzer ⁴⁸	Vocational index ⁴⁸	Interrater reliability (+) <i>Agreement between raters (Cronbach'alpha=.92)</i>	Not assessed
Potvin et al. ⁴⁷	CAPE/PAC ^{5b}	Test-retest reliability (+) <i>Correlation between two independent administration ($r>0.70$ for each subscale except 'CAPE-Social aspect' and 'PAC'; p=not reported)</i>	Content validity (+) <i>Qualitative analysis of the content of the instrument in reference to the construct and the population of interest</i>
Chen et al. ⁴³	Experience sampling method ⁵⁶	Test re-test reliability (?) <i>Consistency in mean scores at two different time points during the week (no differences found at two time points with $p\geq 0.37$ for the T tests in all subscales)</i>	Construct validity (?) <i>Correlation among internal variables</i> <ul style="list-style-type: none"> • (enjoyment with : a) interest $r=0.53$, $p\leq 0.01$; b) anxiety $r=-0.39$, $p\leq 0.01$; • degree of involvement with : a) interest in the situation $r=0.55$, $p\leq 0.01$; • feeling of loneliness with: a) being listened $r=-0.34$, $p\leq 0.01$; b) caring about others $r=-0.33$, $p\leq 0.01$.)
Chen et al. ⁴⁴	Experience sampling method ⁵⁶	Test re-test reliability (?) <i>Consistency in mean scores at two different time points during the week (no differences found at two time points with $p\geq 0.06$ for the T test in all subscales)</i>	Construct validity (?) <i>Correlation among internal variables</i> <ul style="list-style-type: none"> • (enjoyment with : a) interest $r=0.42$, $p\leq 0.01$; • Enjoyment and interest with emotinal ranges, $r=0.20-0.58$, $p\leq 0.05$)

Kramer et al. ^{4b}	PEDI-CAT – responsibility domain ⁶¹	<p>Test-retest reliability (+)</p> <p><i>Intraclass correlation coefficient absolute agreement model (r=0.90; p=not reported)</i></p>	<p>1. Construct validity (+)</p> <p><i>Spearman’s rank correlation coefficient with VABS-II subscales (rho with Communication=0.687, p≤0.001; rho with Daily Living Skills=0.700, p≤0.001; rho with Socialization=0.717, p≤0.001)</i></p> <p>2. Content validity (+)</p> <p><i>Descriptive analysis of the content of the instrument in reference to the population of interest</i></p>
Coster et al. ²⁶	PEDI-CAT – responsibility domain, ASD version ⁵⁸	<p>Internal consistency (+)</p> <p><i>Item Response test analysis such as confirmatory factor analysis (CFI=0.97)</i></p>	<p>1. Construct validity (+)</p> <ul style="list-style-type: none"> • <i>Difference between relevant groups: DIF between the original PEDI-CAT sample and this study sample (DIF for 4 items).</i> • <i>Structural validity: good fit for to a unidimensional model (TLI=0.97; RMSEA=0.05)</i> <p>2. Criterion validity (+)</p> <ul style="list-style-type: none"> • <i>ICC between 5, 10 and 15-items CAT and full item (ICC, 95% CI 5-item=0.95 (0.94–0.96); ICC, 95% CI 10-item=0.98 (0.98–0.99); ICC, 95% CI 5-item=0.99 (0.99–0.99))</i>

<p>McCollum et al.⁴⁶</p>	<p>AYA-ACS⁶⁰</p>	<p>Interrater reliability (?)</p> <p><i>Non-parametric Wilcoxon Signed Rank test between means in the score of each domain completed by each participant and their parents (no differences except for Health and Wellness, $z=-2.66$, $p=0.008$ and trend in Education, $z=-1.9$, $p=0.05$ and Leisure, $z=-1.89$, $p=0.06$)</i></p>	<p>Construct validity (+)</p> <p><i>Spearman Rho correlations between scores in the VABS-II (VABS-II Communication with AYA-ACS Health and Education respectively $\rho=0.57$ and 0.45, $p=0.003$ and 0.03; VABS-II Domestic and internalizing maladaptive behaviour with AYA-ACS Health respectively, $\rho=0.79$ and -0.62, $p=0.000$ and $p=0.01$; VABS-II Community and AYA-ACS Social and Work respectively, $\rho=0.68$ and 0.54, $p=0.004$ and 0.03; VABS-II Leisure with AYA-ACS Chore, $\rho=0.53$, $p=0.03$)</i></p>
<p>Toomey et al.⁴⁹</p>	<p>PROMIS Parent Proxy Peer Relationships Measure⁵⁹</p>	<p>Internal consistency (+)</p> <p><i>Standard error for T-scores (Standard error less than 4 for the T-scores of 92% of the participants, that translates to a reliability coefficient of >0.84)</i></p>	<p>1. Construct validity (+)</p> <ul style="list-style-type: none"> • <i>Correlation with severity of autism as measured by SRS-2 ($r=-0.60$, $p<0.01$)</i> • <i>Differences in PROMIS T-scores between those children who scored severe or moderate in the SRS-2 and those who scored mild or normal in the SRS-2 ($p<0.001$)</i> <p>2. Content validity (-)</p> <p><i>Semi-structured interview about the relevance of the scale for their children (strength reported: questions considered relevant; limitation reported: the answers to items could vary significantly from a week to an other)</i></p>

SFA, school function assessment; ASD, autism spectrum disorder; LD, learning disability; CAPE, children's assessment of participation and enjoyment; PAC, preferences for activities of children; PEDI-CAT: Pediatric Evaluation of Disability Inventory-Computer Adaptive Test; VABS-II, Vineland Adaptive Behaviour Scales, 2nd edition; CFI, comparative fit index; DIF, differential item fit; TLI, Tucker Lewis index; RMSEA, root mean square error of approximation; ICC, intraclass correlation coefficient; AYA-ACS, adolescents and young adult activity card sort; PROMIS, Patient-Reported Outcomes Measurement Information System; SRS-2, Social Responsiveness Scale, Second Edition.

Table III: Main methodological limitations of the studies included and methodological quality

Study	Main methodological limitation/bias (items endorsed in 4-points COSMIN checklist)	Methodological quality (poor/fair/good)
Davies et al. ⁵⁰	<ul style="list-style-type: none"> • Small sample size ($n=11$) • % of missing items not described. 	poor
Potvin et al. ^{47,a}	<ul style="list-style-type: none"> • Small sample size ($n=14$ and $n^a=27$) • % of missing items not described 	
Chen et al. ⁴³	<ul style="list-style-type: none"> • Small sample size ($n=4$) • doubtful if time interval (test-retest reliability) was appropriate • not clear reason below the statistical method used 	
Chen et al. ⁴⁴	<ul style="list-style-type: none"> • Small sample size ($n=6$) • doubtful if time interval (test-retest reliability) was appropriate • not clear the reason below the statistical method used 	
Kramer et al. ⁴⁵	<ul style="list-style-type: none"> • Small sample size ($n=29$; $n^a=34$) • % of missing items not described 	
McCullum et al. ⁴⁶	<ul style="list-style-type: none"> • Small sample size ($n=24$) • % of missing items not described 	
Taylor and Seltzer ⁴⁸	<ul style="list-style-type: none"> • Moderate sample size ($n=34$) • % of missing items not described 	fair
Toomey et al. ⁴⁹	<ul style="list-style-type: none"> • % of missing items not described. 	good
Coster et al. ²⁶	<ul style="list-style-type: none"> • % of missing items not described 	

This table describes the main methodological limitations for each study included in this systematic review. We used the 4-points COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist³⁴ to determine the rating of the methodological quality for each study and this is described in the last column of this table. The second column reports the main methodological limitation that two independent reviewers endorsed for each study. For each study, the first item marked in bold represents the limitation that determined the overall quality for that study. ^aSample size for the study of validity.

Table IV: Levels of evidence for the quality of the measurement property for each instrument

Tool	Study	Rating of the measurement property, methodological quality of the study		Level of evidence for each tool	Domain(s) of participation	Dimension(s) of participation
		Reliability	Validity			
SFA ⁵⁴	Davies et al. ⁵⁰	<i>not assessed</i>	+, poor methodological quality	only a study of poor methodological quality	educational activities	extent of involvement
Vocational index ⁴⁸	Taylor and Seltzer ⁴⁸	+, fair methodological quality	<i>not assessed</i>	limited	vocational activities and employment	extent of involvement
CAPE/PAC ⁴⁷	Potvin et al. ⁴⁷	+ CAPE - PAC, poor methodological quality	+, poor methodological quality	only a study of poor methodological quality	domestic life (household task), educational activities and employment leisure activities (recreation, play, hobbies, sport)	extent of involvement enjoyment of participation
ESM ⁵⁶	Chen et al. ^{43,44}	?, poor methodological quality	?, poor methodological quality	only studies of poor methodological	domestic life (household task), interpersonal	extent of involvement

Tool	Study	Rating of the measurement property, methodological quality of the study		Level of evidence for	Domain(s) of participation	Dimension(s) of participation
				quality	relationship, leisure activities, educational activities	enjoyment of participation
PEDI-CAT ⁶¹	Kramer et al. ⁴⁵	+, poor methodological quality	+, fair methodological quality	limited (for validity only)	extent of participation in daily task	extent of involvement
PEDI-CAT, ASD version ⁵⁸	Coster et al. ²⁶	+, good methodological quality	+, good methodological quality	moderate	extent of participation in daily task	extent of involvement
AYA-ACS ⁶⁰	McCollum et al. ⁴⁶	?, poor methodological quality	+, poor methodological quality	only a study of poor methodological quality	domestic life (household task), caring for others, interpersonal relationship, leisure activities, educational activities, vocational training and employment	extent of involvement enjoyment of participation
PROMIS	Toomey et al. ⁴⁹	+,	+ -,	limited	interpersonal	extent of

Tool	Study	Rating of the measurement property, methodological quality of the study		Level of evidence for	Domain(s) of participation	Dimension(s) of participation
Parent Proxy Peer Relationships Measure ⁵⁹		good methodological quality	good methodological quality	(for reliability and construct validity only)	relationships, community life and caring for others	involvement

This table summarizes the findings of the measurement properties (+=positive rating; -=negative rating; ?=indeterminate rating) and the methodological quality of the included studies by providing an overall level of evidence for each tool included. The evidence was rated following the recommendation of Terwee⁴²: moderate, if there was only one study of good methodological quality; limited, if there was only one study of fair methodological quality; indeterminate, if there was/were only study/ies of poor methodological quality. In the last two columns, this table also reports a summary about the domain and dimensions of participation investigated by each tool. SFA, school function assessment; CAPE, children's assessment of participation and enjoyment; PAC, preference for activities of children; ESM, experience sampling method; PEDI-CAT, Pediatric Evaluation of Disability Inventory, Computer Adaptive Test; PEDI-CAT, ASD, Pediatric Evaluation of Disability Inventory, Computer Adaptive Test, autism spectrum disorder version; AYA-ACS, adolescents and young adults activity card sort.

Author Manuscript

Author Manuscript

Figure 1: Flow chart of the study inclusion