



Published in final edited form as:

J Autism Dev Disord. 2024 July ; 54(7): 2625–2635. doi:10.1007/s10803-023-05945-4.

Item Understanding of Common Quality of Life Measures for use with Autistic Adults

Kelly B. Beck¹, Lauren A. Terhorst², Carol M. Greco¹, Jamie L. Kulzer³, Elizabeth R. Skidmore², Michael P. McCue³

¹ Department of Psychiatry, University of Pittsburgh School of Medicine, 303 Webster Hall 4415 Fifth Avenue, Pittsburgh, PA 15213, USA

² Department of Occupational Therapy, University of Pittsburgh School of Health and Rehabilitation Sciences, Pittsburgh, USA

³ Department of Rehabilitation Science and Technology, University of Pittsburgh School of Health and Rehabilitation Sciences, Pittsburgh, USA

Abstract

Quality of life (QOL) and life satisfaction are important research priorities for autistic adults. As such, we saw a need to evaluate individual items of commonly used subjective QOL scales to understand how they are interpreted and perceived by autistic adults. This study used cognitive interviews and repeated sampling to evaluate the accessibility, test-retest reliability and internal consistency of several common QOL measures in a sample of young autistic adults ($n = 20$; aged 19–32). Cognitive interviews suggested that the Satisfaction with Life Scale was well understood and demonstrated excellent internal consistency and test-retest reliability. While the WHOQoL-BREF and WHOQoL Disability Modules had adequate reliability, cognitive interviews suggested that additional instructions and examples would further enhance their accessibility for use with autistic adults.

Keywords

Autism spectrum disorder; Quality of life; Adults; Accessibility; Cognitive interviews; Patient-reported outcome measures

Quality of life (QOL) and life satisfaction outcomes are high research priorities for the autistic community (Benevides et al., 2020; Ne’eman, 2010; Robertson, 2010). In a recent study, QOL was ranked the number one outcome that mattered to autistic adults (Benevides et al., 2020). Both subjective quality of life and life satisfaction are indicators of overall wellbeing; though there is a lack of clarity on the distinction between the two constructs. Quality of life is defined by the World Health Organization (WHO) as, “an individual’s perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns” (World Health Organization. Division of Mental Health, 1996). Global life satisfaction has been defined as

[✉] Kelly B. Beck kellybeck@pitt.edu.

“evaluation of one’s life” (Diener et al., 1985; Diener, 1984). Subjective QOL is typically measured by assessing satisfaction across several key life domains while life satisfaction is a global evaluation. Both involve an individual’s subjective and cognitive assessment of their lives. Considering the priority for autistic people and the subjective nature of the constructs, precise and accessible self-report instruments of QOL and life satisfaction are essential.

It is widely documented that subjective QOL is lower for autistic people than nonautistic people (van Heijst & Geurts, 2015). However, previous QOL measurement may have been imprecise with frequent use of proxy measurement and measures not developed or validated for autistic individuals (van Heijst & Geurts, 2015). Subjective QOL measures initially developed for neurotypical people may be interpreted differently by autistic people, posing a methodological barrier to accurately capturing QOL within the autism community. For instance, QOL measures may utilize a set of neuronormative domains, life domains selected and viewed as important to neurotypical people, such as physical health, psychological health, and social relationships (Lam et al., 2021; McConachie et al., 2020; Nicolaidis et al., 2020; Pellicano & den Houting, 2022; Robertson, 2010). Domains and items prioritized in a neurotypical-majority society could be interpreted or valued differently by an autistic person, risking measurement error of this subjective construct. As such, researchers must consider QOL and life satisfaction measurement from the perspective of autistic people.

The lack of appropriate measures in autistic adult research is not unique to assessing QOL and life satisfaction constructs. Recent community-engaged work concluded that many self-report outcome measures have common challenges for use with autistic adults, such as problematic language complexity and pragmatics, imprecise response options, problematic time references in the instructions, and construct validity (Nicolaidis et al., 2020). These authors note that some of these challenges are related to language and sentence structure complexity on measures but also confusion or concrete interpretation of common figures of speech (Nicolaidis et al., 2020). Authors recommend visual scales to supplement Likert response scales and highlight that autistic people often feel that scales developed with neurotypical samples do not capture their experience (Nicolaidis et al., 2020). It is, therefore, important that researchers evaluate the accessibility and understanding of commonly used self-report QOL and satisfaction measures with autistic adults (Nicolaidis et al., 2020).

The WHOQoL-BREF is one commonly used outcome measure in adult autism research (McConachie et al., 2018; van Heijst & Geurts, 2015). The WHOQoL-BREF is a brief QOL measure developed by the WHO to be culturally valid across over 20 countries worldwide (Skevington et al., 2004; World Health Organization, 1996). The WHOQoL then developed a WHOQoL-Disabilities Module (WHOQoL-DIS) for individuals with physical or intellectual disabilities, which also included an adapted version of the WHOQoL-BREF for individuals with intellectual disability (Power et al., 2010; The WHOQOL-Dis Group et al., 2010). However, autistic individuals were not included in the development or validation of these measures. The Satisfaction with Life Scale (SWLS) is another widely used measure in general adult literature, yet recent work has similarly highlighted the lack of autistic perspectives in measuring the construct of satisfaction and wellbeing (Lam et al., 2021).

Notably, McConachie et al. addressed one major methodological barrier, construct validity, in QOL measurement in autism. McConachie et al., consulted with the autistic community to develop 9 autism specific QOL items (ASQoL) to add to the WHOQoL-BREF and WHOQoL-DIS (McConachie et al., 2018). This group also evaluated the psychometric properties of the WHOQoL-BREF, WHOQoL-DIS, and the ASQoL in an autistic adult sample (McConachie et al., 2018). Even with these major advances, however, the individual items on the WHOQoL-BREF, WHOQoL-DIS, and SWLS have not been evaluated for accessibility and understanding in autistic adults. It is important to evaluate how autistic adults interpret and understand individual items on these scales given that the measures were originally developed in neurotypical samples.

The aims of this study were to test that the WHOQoL-BREF, WHOQoL-DIS and SWLS measures are well understood, demonstrate internal consistency ($\alpha \geq 0.8$), and adequate test-retest reliability ($ICC \geq 0.75$) among autistic adults. To test these aims we conducted cognitive interviews and repeated sampling reliability statistics on versions of the WHOQoL-BREF, WHOQoL-DIS, and SWLS in a sample of young autistic adults living in a postsecondary education facility ($n = 20$). We first evaluated the WHOQoL-BREF, WHOQoL-DIS, and SWLS using cognitive interviews and a test-retest design. Participants provided feedback that the WHOQoL and WHOQoL-DIS were difficult to understand. In response to these findings, we re-enrolled participants to evaluate the WHOQoL versions developed for intellectual disabilities (WHOQoL-BREF-ID and WHO Disabilities-ID) six months later.

Method

In a cohort, prospective design, autistic adult research participants ($n = 20$) completed a brief QOL assessment battery, including the WHOQoL-BREF, WHOQoL-DIS, and SWLS, at two timepoints with a washout period of two weeks (Shoukri et al., 2004). A subset of participants ($n = 8$) completed cognitive interviews on each measure to evaluate the understanding of each item on the assessment. A ‘cognitive interview’ is an open-ended interview to assess participant comprehension and readability of each individual item and response options in a measure or item pool (Beatty & Willis, 2007; *PROMIS*, 2013). Cognitive interview participants provided feedback that several items on the WHOQoL-BREF and WHOQoL-DIS were challenging to understand. In an interest to learn more, we offered participants the option to re-consent and re-enroll in the study 6 months later to evaluate the WHOQoL-BREF-ID and WHOQoL-DIS-ID scales using the same methods (i.e. test-retest with two week washout in entire sample and cognitive interviews with subset of 8 participants). Two participants from the original cohort were not available for the evaluation of ID versions, and, in response, two new participants enrolled.

Sample

The total sample consisted of 22 young autistic adults. A recent methodological review indicated that a small sample is needed for test-retest reliability when the goal is to show agreement of a participant’s score on a scale between two timepoints (Bujang & Baharum, 2017). As the level of agreement increases, the sample size needed decreases. For example, a

sample of approximately 20 participants will be able to detect an ICC when agreement is as low as 0.5, and a sample as small as 10 participants can detect an ICC of 0.7.

This sample of young autistic adults were recruited from a postsecondary vocational training facility. This facility provides comprehensive rehabilitation services and educational training for over 300 students with disabilities in one building. Participants live at the facility and are enrolled in a variety of post-secondary vocational training programs or trade school programs. Associate degrees and certificate programs are the degree options at this facility. Participants were recruited through flier advertisements and announcements in classrooms and social groups. Upon consenting to enroll in the study, participants' public vocational rehabilitation case files were reviewed to ensure basic eligibility criteria.

Participants met inclusion criteria if the case file had a documented disability diagnosis of Autism Spectrum Disorder, Pervasive Developmental Disorder, or Asperger's Disorder. Participants were excluded from the study if: (1) they had a mental health counseling session scheduled during the 2-week test-retest assessment period to preclude any changes in QOL measures due to intervention-related changes; (2) case files included an Intellectual Disability diagnosis or documented full scale IQ score < 70 [measured by the Wechsler Adult Intelligence Scale (WAIS) III, WAIS IV, Multidimensional Aptitude Battery (MAB), or Kaufman Brief Intelligence Test (K-BIT)]; (3) case files indicated verbal intelligence scores < 75 (measured by the WAIS III, WAIS IV, MAB, or K-BIT); (4) case files included WAIS III or WAIS IV similarities subscale score < 5; or (5) the participant was not interested in participating in the study.

The mean age of participants was 20.7 ($SD = 2.98$) with a range of 19–32. The sample consisted of 81.8% males and 18.2% females (sex assigned at birth), with 95.5% of participants self-identified as Caucasian and 4.5% as Native American. All participants successfully completed high school with a high school diploma or GED. All participants had a primary diagnosis of Autism Spectrum Disorder, Pervasive Developmental Disability, or Asperger's. Co-occurring diagnoses included the following: Attention Deficit Hyperactivity Disorder (63.6%), Depressive Disorder (31.8%), Generalized Anxiety Disorder (13.6%), Mood Disorder NOS (9%), Obsessive Compulsive Disorder (4.5%), seizures (4.5%), and hearing loss (4.5%). Participants were able to participate in the study while still receiving active mental health services (e.g. psychopharmacology, outpatient services, supportive counseling at the postsecondary education facility). In order to avoid any confounding impact of intervention on QOL assessment scores, the test-retest study appointments were scheduled during a 2-week time period when participants did not have scheduled mental health service appointments.

The full-scale IQ sample mean was 94.35 ($SD = 12.69$) with a range of 80–133. The mean Verbal Comprehension Index IQ scores of participants was 101.75 ($SD = 16.9$) with a range of 83–147. When available in case files ($n = 11$), similarities subtest scores ranged from 7 to 17, with a mean of 10.5 ($SD = 3.17$).

Procedure

All participants provided written consent to enroll in this study, which was approved by the University of Pittsburgh Institutional Review Board (PRO16050352). Participants independently completed an assessment battery consisting of the SWLS, WHOQoL -BREF, WHOQoL -DIS at two time points (baseline and 2 weeks). As noted above, the WHOQoL-BREF-ID and WHOQoL-DIS-ID procedures were completed 6-months after the initial assessment battery using the same strategy (two time points; baseline and 2 weeks). The same research assistants administered the assessments to all participants at all timepoints to ensure standardization. Research assistants were available for participants that had questions or needed support completing the questionnaires. Research assistants were instructed to first read the question aloud that was causing difficulty and then clarify confusing phrasing if the participant was still having difficulty. Research assistants documented any instances of providing support on the questionnaires. Participants were given the choice to complete the assessment battery online via Qualtrics or on paper. Following completion of the second assessment battery, participants received \$10 compensation.

Trained research staff conducted in person cognitive interviews with a subset (n = 8) of enrolled participants at time point 1 in the test-retest protocol. Cognitive interviews are a structured method used in measure development to ensure understanding and readability of individual items in a patient reported outcome scale (Beatty & Willis, 2007; PROMIS[®], 2013). All cognitive interview participants completed the cognitive interview process with every measure and all items on each measure. This exceeds guidelines suggesting that at least 5 participants evaluate each individual item of a measure or item pool (PROMIS[®], 2013).

Interviewers used a structured document to guide the cognitive interview process, and interviews were audio recorded and transcribed. Cognitive interviews were conducted using a combined think-aloud and debriefing methodology. Specifically, participants were asked to read each item of all measures aloud, reword the question into their own words, and think aloud as they chose their answer. Participants were also asked to describe the directions, response sets, and items in their own words aloud and provide feedback. Interviewers used open-ended questions, avoided summarizing participant responses, and refrained from using reflections to reduce bias and limit influencing participant responses. Participants were asked to circle any confusing item wording and explain the reasoning for selecting circled words. Interviewers took notes on any responses that suggested confusion or lack of clarity with the item.

Immediately after the cognitive interviews for each scale, cognitive interview participants then answered two questions about their global understanding of the scale. The items included, “the questions were easy to understand” and “the wording of the questions made sense to me.” Participants selected responses on these two questions using a 4-point response option (0 = Not at all, 1 = a little bit, 2 = quite a bit, 3 = very much). These questions were used to quantify agreement of understanding for each scale. In addition, interviewers documented impressions after each cognitive interview to identify patterns across items that would not be reflected in the item-level notes.

Upon completion of the cognitive interviews, two study investigators met to discuss participant feedback and data. Interview notes and 4-point scale understanding questions were considered alongside item-level feedback (e.g. circled phrases, read aloud comments). Consensus was made to identify items on each scale that were problematic for participant understanding.

Measures

Satisfaction with Life Scale (SWLS)

The SWLS is a global measure of satisfaction with QOL and well-being (Diener et al., 1985). The SWLS is a five-item measure of holistic well-being and uses a 7-point Likert rating scale. This scale reports a reliability alpha coefficient of 0.87 at the time of the assessment and an alpha coefficient of 0.82 two months following the original assessment. This measure was initially validated in young adults (undergraduate students) and older adult populations.

WHOQoL-BREF

The WHOQoL-BREF is a 26-item internationally developed global QOL measure of general well-being (World Health Organization, 1996). This scale has acceptable reliability alpha coefficients across four subscale domains ($\alpha = 0.82$, physical health; $\alpha = 0.81$, psychological; $\alpha = 0.80$, environment; $\alpha = 0.68$, social relationships (Skevington et al., 2004). This measure was developed with neurotypical adults aged 12–97 ($M = 45$, $SD = 16$).

WHOQoL-BREF-ID

A simplified version of the WHOQoL-BREF was developed for individuals diagnosed with Intellectual Disabilities, WHOQoL -BREF-ID (Power et al., 2010; The WHOQOL-Dis Group et al., 2010). This version contains the same root items but with simpler phrasing and a 3 or 5-point visual response scale (Fang et al., 2011; Power et al., 2010). This version has acceptable internal consistency for each subscale ($\alpha = 0.705$, physical health; $\alpha = 0.759$, psychological; $\alpha = 0.792$, environment; $\alpha = 0.606$, social relationships). This scale was developed in a sample of individuals with physical ($n = 909$) or intellectual ($n = 491$) disabilities, and this scale used with permission from originating authors (Power et al., 2010).

WHOQOL-DIS

International focus groups suggested that the WHOQoL was missing several important QOL factors for individuals with either physical or intellectual disabilities (Power et al., 2010). Thus, a 13-item disability-specific QOL measure was developed, the WHOQoL-DIS, and has a 3-factor structure, (1) discrimination, (2) autonomy, and (3) inclusion. The WHOQoL-DIS can be administered with either a 3-point or 5-point Likert visual response scale and has previously demonstrated adequate internal consistency ($\alpha = 0.852$). This scale was developed in a sample of individuals with physical ($n = 909$) or intellectual ($n = 491$) disabilities, and this scale used with permission from originating authors (Power et al., 2010).

WHOQOL-DIS-ID

A simplified version, the WHOQoL -DIS-ID, uses a 3 or 5-point visual response scale was used for individuals with intellectual disabilities (Fang et al., 2011). The WHOQoL -DIS-ID has acceptable internal consistency ($\alpha = 0.808$). The WHOQoL-DIS-ID can be administered with either a 3-point or 5-point Likert visual response scale. The 5-point visual response scale was used for all test-retest reliability analyses. Consistent with cognitive interview procedures to review all directions and response options, cognitive interview participants were shown both the 3-point and 5-point response and asked for any feedback or preference. This scale was used with permission from originating authors (Power et al., 2010).

Analyses

The following three research questions guided the analyses: (1) Does each instrument demonstrate internal consistency ($\alpha > 0.8$)? (2) Does each instrument demonstrate test-retest reliability ($ICC > 0.8$)? (3) Is each instrument well understood (measured in cognitive interview data and agreement on two understanding items)? Internal consistency reliability estimates were computed using Cronbach's α (Gliner et al., 2001). Alpha levels greater than 0.8 were acceptable values. Test-retest reliability was evaluated with interclass correlation coefficients (ICC), Eq. 2. The following guidelines were utilized to interpret ICC : <0.5 poor; $0.5-0.75$ moderate; $0.75-0.9$ good, and > 0.9 excellent (Koo & Li, 2016).

Understanding was evaluated by interpreting all cognitive interview data and agreement scores from the two understanding questions. Scores on the understanding questions ("the questions were easy to understand" and "the wording of the questions made sense to me." with the response scale 0 = Not at all, 1 = a little bit, 2 = quite a bit, 3 = very much) were dichotomized into agreement or disagreement. Scores of a 2 or 3 constituted agreement. Scores of 0 or 1 constituted disagreement. Percentages of agreement were computed for each measure. Two investigators synthesized agreement ratings with cognitive interview data (e.g. circled items of confusion, think aloud comment, interviewer notes, and transcripts). Analyses identified measures and specific items within measures that were causes confusion. These findings were classified by type of challenge (e.g. lexical/phrasing, temporal) and when the challenge occurred (i.e. understanding item, choosing response). Statistical analyses were run using SPSS, Version 22.

Results

The SWLS met a priori benchmarks acceptable internal consistency, excellent test-retest reliability, and was well understood. Participants had difficulty interpreting items on the WHOQoL-BREF and WHOQoL-DIS, which was reflected in the cognitive interviews and agreement scores from understanding questions. The ID versions of the WHOQoL-BREF and WHOQoL-DIS had better interpretation, however, the social relationship domain items still caused some confusion among participants. Table 1 summarizes agreement and reliability statistics per measure. Table 2 details item-level challenges and classifications on all measures.

Satisfaction with Life Scale (SWLS)

Cognitive Interviews

Cognitive data review indicated that participants interpreted and answered the questions without requiring assistance or questioning from interviewers. Participants' answers reflected appropriate understanding of each item. Participant responses to the understanding questions reached 100% agreement.

Reliability Analyses

The SWLS demonstrated acceptable internal consistency ($\alpha = 0.789$) and excellent test-retest reliability ($ICC_2 = 0.948$). These results suggest that the SWLS is an accessible life satisfaction measure for autistic adults.

WHOQoL-BREF

Cognitive Interviews

A review of the cognitive interview transcripts illuminated several challenges, which were classified into difficulty understanding the items and choosing responses (Table 2). At the item-level, four items were the most problematic and caused challenges due to unclear phrasing of the item (lexical) and confusion with the item concept or construct. Further, more than half of the participants demonstrated challenges with the temporal guidance in the scale instructions. Specifically, five of the eight cognitive interview participants frequently described rationale of answer selections that were outside of the 2-week reference. In addition, interviewers noted that items that referred to living place, local activities, and physical environment were answered inconsistently. Participants were sometimes answered these questions about 'home' and then would change to answer by referencing their current setting (residing at a postsecondary education dorm). Participants reached 75% agreement on the two understanding questions for this measure.

Three of the eight cognitive interview participants found the phrase "to what extent" in item 3 ("To what extent do you feel that physical pain prevents you from doing what you need to do?") to be difficult. Four of the eight cognitive interview participants required assistance answering item 4 ("How much do you need any medical treatment to function in your daily life?"). Participants were not sure what constituted medical treatment and questioned if this included daily medications. Six of the eight cognitive interview participants requested assistance for item 21 ("How satisfied are you with your sex life?"). Item 21 was difficult to answer for those participants who were not sexually active. The participants found it confusing to answer on their *satisfaction* with their sexual inactivity. Finally, over half of the cognitive interview participants had difficulty with the word 'transport' in item 25 ("How satisfied are you with your transport?"), which is likely due to cultural differences in this scale designed for international use. Overall, item 21 ("How satisfied are you with your sex life?") caused the most confusion amongst participants.

Reliability Analyses

The WHOQoL-BREF demonstrated excellent test-retest reliability across all domains, ranging from $ICC_2 = 0.83 - 0.95$. However, the WHOQoL-BREF demonstrated variable internal consistency across domains, ranging from $\alpha = 0.387 - 0.851$. Item 4 (“How much do you need any medical treatment to function in your daily life?”) was eliminated from the physical health domain due to qualitative evidence from the cognitive interviews and internal consistency was improved from $\alpha = 0.387$ to $\alpha = 0.728$. Item 22 (“How satisfied are you with the support you get from your friends?”) contributed to problematic internal consistency. Internal consistency improved from $\alpha = 0.548$ to $\alpha = 0.649$ when this item was eliminated. There are only three items in the social relationships domain, and two of the items indicated problems in this study sample (item 21 ‘sex life’ and item 22 ‘support from friends’).

WHOQOL-BREF-ID

Cognitive Interviews

Participants had less difficulty completing the WHOQOL-BREF-ID version. The WHOQoL-BREF-ID received a 100% agreement rating across participants on the two understanding questions. However, interviewer notes indicated that participants still explained answers outside of the two-week time reference point on items related to living place, local activities, and physical environment. Participants were asked to think aloud on the two difference response options on this scale (5-point or 3-point response scales; Fang et al., 2011). All 8 participants indicated a preference for a 5-point response scale over the 3-point response scale (Fang et al., 2011).

Item-level analyses of the cognitive interviews indicated that only one item remained a challenge for participants. Five participants still required assistance from the interviewer on item 21 (“Are you satisfied with your sex life, or your relationship with your partner?”; Table 1). Many participants still indicated that they could not answer the question because they were not sexually active or in a relationship. The interviewer asked participants to consider their satisfaction with not having an active sex life or relationship. The ID version corrected confusion on items 3, 4, and 25 with the changes in phrasing and examples provided with each question.

Reliability Analyses

The WHOQoL-BREF-ID demonstrated excellent test-retest reliability across all domains, ranging from $ICC_2 = 0.87 - 0.93$. However, the WHOQoL-BREF-ID still demonstrated variable internal consistency ranging from $\alpha = 0.464 - 0.821$ across domains. Only the psychological domain demonstrated acceptable internal consistency ($\alpha = 0.821$). Item 4 (“Do you need any medical treatment to help you in your daily life? *For example, medicines*”) performed similarly in the ID version as the original WHOQoL-BREF. Once again, internal consistency was improved with the elimination of item 4 from $\alpha = 0.464$ to $\alpha = 0.704$. The social relationships domain still did not perform well in qualitative or quantitative analyses. Internal consistency was improved with the elimination of item 22 from $\alpha = 0.558$ to $\alpha = 0.767$ (“Are you satisfied with the support you get from your

friends?"). Alpha levels for the environmental domain were lower than acceptable standards ($\alpha = 0.696$).

WHOQoL-DIS

Cognitive Interviews

Similar to the other measures, interviewer notes from cognitive interviews indicated that several participants explained answers that were outside of the two-week time reference of the scale. Item-level data suggested that participants had some problems with the phrasing in the WHOQOL-DIS module. One participant had difficulty answering the item 28, "do you feel that some people treat you unfairly?" The participant indicated they were unsure what 'some' meant and wanted more concrete instructions. This impacted their response choice. Another participant indicated that the example included in item 30 was problematic, "Do you worry about what might happen to you in the future? *For Example, thinking about not being able to look after yourself or being a burden to others in the future*". This participant was unsure of the meaning of 'burden' and asked to cross out the example. This impacted the understanding of the item, but the participant was able to select a response and explain their choice once the example was crossed out. Participants reached 75% agreement on the understanding questions for this measure.

Reliability Analyses

The WHOQoL-DIS demonstrated excellent test-retest reliability across all factor domains and total score, ranging from $ICC_2 = 0.76 - 0.88$. The WHOQoL-DIS also demonstrated variable internal consistency across factor domains, ranging from $\alpha = 0.352 - 0.881$ across domains. The discrimination domain (3 items) had the lowest internal consistency ($\alpha = 0.352$), but it did not improve with the elimination of any items.

WHOQOL-DIS-ID

Cognitive Interviews

Cognitive interview data and agreement rating (100%) indicated that participants did not have difficulty with any items on WHOQOL-DIS-ID module. Notably, the problematic phrases identified in the WHOQoL-DIS were not revised in the ID version, yet items 28 and 30 were not identified again in these interviews. The participant that initially identified WHOQoL-DIS items 28 and 30 as confusing did not re-enroll to evaluate the ID version. Thus, we were unable to interpret if this was due to familiarity with the scale or differences in perspectives across participants. Interviewer notes indicated that several participants explained answers that were outside of the two-week time reference.

Reliability Analyses

The WHOQOL-DIS-ID demonstrated adequate test-retest reliability across all factor domains and total score, ranging from $ICC_2 = 0.786 - 0.942$. The WHOQoL-DIS-ID also demonstrated variable internal consistency across factor domains, ranging from $\alpha = 0.412 - 0.924$ across domains. The discrimination domain (3 items) had the lowest internal consistency ($\alpha = 0.412$), but improved with the elimination of item 30 ($\alpha = 0.612$; Do you

worry about what might happen to you in the future? *For example, thinking about not being able to look after yourself, or being a burden to others in the future*). The autonomy (3 items) domain demonstrated an alpha of 0.689 but improved with the elimination of item 33 ($\alpha = 0.707$; “Do you get to make the big decisions in your life? *For example, like deciding where to live, or who to live with, how to spend your money*”). Although internal consistency improved with the elimination of items, the cognitive interviews did not indicate confusion or difficulty with these items. Thus, they were not eliminated in final analyses. See Table 1 below for agreement and reliability statistics.

Discussion

This study suggests that young autistic adults can reliably and consistently report on QOL and life satisfaction, as all selected measures demonstrated excellent agreement in the test-retest analyses. Importantly, the SWLS appears to be an accessible measure of life satisfaction for young autistic adults. This is the first study, to our knowledge, that examined the accessibility and reliability of the SWLS in an autistic sample. Participants understood each item in the cognitive interviews and the scale demonstrated strong internal consistency and test-retest reliability in this sample. Further, the SWLS is low burden to participants with only 5 items that are easily understood. This study did not, however, evaluate the construct validity of the SWLS to determine if there are items missing to fully capture life satisfaction from the perspective of autistic adults. Future work may involve using mixed methods to further explore the construct validity and measurement of this important outcome (Lam et al., 2021; Nicolaidis et al., 2020).

WHOQoL internal consistency and test-retest reliability statistics in this study were fairly consistent with larger studies examining the psychometric properties of the WHOQoL-BREF in autistic adult samples (McConachie et al., 2018). Internal consistency statistics in the physical health subdomain of the WHOQoL-BREF and the discrimination subdomain of the WHOQoL-DIS were lower in this sample compared to McConachie’s work (McConachie et al., 2018). This was improved in our sample when removing the confusing medication item from the physical health domain.

The ID versions of the WHOQOL-BREF and WHOQoL-DIS caused less confusion among autistic adults in this sample. The revisions made for the ID versions, particularly of the WHOQoL-BREF, eliminated confusing phrases of speech, added examples to each question, and added visuals to the response options, all of which are consistent with guidelines for creating accessible survey instruments for autistic adults (Nicolaidis et al., 2020). Thus, it is reasonable for researchers to consider using the ID versions of the WHOQoL with the 5-tem response option. The items that caused confusion on the WHOQoL-DIS were not changed in the ID version, though participants did not identify them as confusing in the WHOQoL-DIS-ID cognitive interviews. This may be due to the initial participant not re-enrolling to evaluate the ID versions. Despite this discrepancy, the ID versions are still likely advantageous due to the added examples, simplified language, and visual response options. Further, the WHOQoL-BREF and WHOQoL-DIS are designed to be administered together with the compatible versions.

The WHOQoL social relationship domain posed the most challenges in interviews and reliability analyses. Many participants had difficulty rating their satisfaction with sexual relationships when it did not apply to them, which is likely due to concrete interpretation of the phrasing in the item. The ID version of this question did not seem to help the confusion, as it states, “Are you satisfied with your sex life, or your relationship with your partner? *For example, your husband/wife, boyfriend/girlfriend.*” Participants still had difficulty with this question despite the added example in the ID version and previously completing this scale 6-months prior. Researchers might consider adding a different example that clarifies how to answer if the participant is not sexually active in addition to revising the example to revising for gender neutral wording. Given the confusion with this item, it is not surprising that the internal consistency was problematic on this subdomain. Adding more relevant items within the social relationship subdomain could also potentially improve reliability. Notably, the ASQoL items designed to be added to the WHOQoL-BREF do include one additional social relationship item, “Are you satisfied with your current friendships? (i.e. whether you have several, few, or no friends)”, which may be a helpful item for these efforts (McConachie et al., 2018). Future work may consider further exploring what QOL means in relation to social relationships, with autistic people guiding this work. It is possible that this whole domain is being considered from a neurotypical perspective and additional aspects are not being assessed in any available measures.

Cognitive interviews suggested that many participants had challenges with the time reference and any item that included descriptors of place of residence. Participants recalled instances substantially (sometimes several years) outside of the 2-week time reference when considering their response, which is something researchers should consider when using these WHOQoL tools. Interviewer notes suggest that participants were recalling very salient negative past events, which may still influence current evaluation of QOL as much as recent events. This is consistent with other work finding that instances of bullying predicted QOL on the WHOQoL-BREF (Hong et al., 2016). Though, interviewers did not ask participants to explain why their reasoning for selecting answers was outside of the 2-week time period (per protocol to avoid influencing responses). Thus, it is unclear if this is related to QOL stability or the significance of past events on current evaluation of subjective life satisfaction. Participants also inconsistently responded to items related to place of residence. The same participant would sometimes reference their hometown on these items and then reference their current living arrangements on the postsecondary education campus on other items. Researchers conducting studies in transition-age autistic adults temporarily living away from home might consider including examples and instructions specifying which residence to consider when completing the survey.

The SWLS, WHOQoL-BREF, and WHOQoL-DIS have some minor overlap between the measures. Overall, the WHOQoL measures assess specific life domains while the SWLS includes 5 items on global life satisfaction (e.g. *in most ways my life is close to my ideal; the conditions of my life are excellent; I am satisfied with my life; so far I have gotten important things I want in life; if I could live my life over; I would change almost nothing*). The WHOQoL-BREF-ID does include two global items, “*How would you rate your quality of life?*” and “*How satisfied are you with your health?*”. The WHOQoL-BREF global items differ slightly from life satisfaction and are not included in the module scoring for the scale.

This study did not assess the relationship between life satisfaction and QOL scores with other constructs, outcomes, and characteristics. Thus, findings from this study are not able to speak to the value of administering both the WHOQoL (multidomain) and SWLS (global evaluation) measures together.

This study was preliminary and a first step to assessing the understanding of commonly used QOL and life satisfaction measures in autistic adult research. As such, this study had methodological limitations that impact the generalizability of these findings. This sample was limited to young adults with a GED or high school diploma, predominantly male, without co-occurring intellectual disability and living away from home with supports from the facility. While nearly 60% of the sample had co-occurring mental health diagnoses, this convenience sample was limited to a narrow age range of transition-age young adults receiving services at one facility. Further, participants were recruited from a state rehabilitation vocational facility where diagnoses are made and tracked cumulatively from medical, psychiatric, and high school records. Thus, the disability diagnoses of this study population are not those of a clinical research study and no data on socioeconomic status was available for collection. While these results can possibly generalize to other young adults on the spectrum in other postsecondary education settings, they may not be applicable to older adults living in the community or those with intellectual disability. These sample limitations likely directly impacted the interpretation of several WHOQoL items, given that living in a temporary educational dorm setting contributed to confusion on questions with location descriptors.

Despite these limitations, this study suggests that accessibility and understanding of a commonly used QOL measure can be improved. This warrants future exploration and testing with larger samples that more reflective of the entire spectrum and lifespan. Research in nonautistic populations have highlighted QOL disparities among historically excluded groups (Bukavina et al., 2017); thus, there is a need to consider issues of intersectionality in autistic QOL research with accessible and sensitive measures in more representative samples. Future work might also consider more exploration into the time references on self-report measures, both individually and in conjunction with other scales. Self-report measures are often deployed in large batteries. Together, differences in instructions, response options, and time references will likely impact the understanding and responses on these measures and poses another methodological barrier to autistic adult research.

Conclusion

This study contributes to the advancement and understanding of QOL measurement in autism. Cognitive interviews in this study suggest that additional instructions and examples are needed for several items when administering the WHOQoL in autistic adult research samples, which is consistent for emerging guidance on self-report survey research (Nicolaidis et al., 2020). These examples and additional instructions should be developed directly with autistic adults. The SWLS offers a low-burden, reliable, and easily understood option for global life satisfaction. Researchers and practitioners interested in more specific domains may consider utilizing the WHOQoL-BREF-ID with additional instructions and examples for the medication and social domain questions in conjunction with the ASQoL

module (McConachie et al., 2018). Future work should further explore the construct validity of the SWLS to measure life satisfaction in autistic adults. Together, these measures may further advance research in directions that the community wants and supports.

Acknowledgements

This research was funded, in part, by the University of Pittsburgh School of Health and Rehabilitation Sciences Research Development Fund. This work was also supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number KL2TR001856 (Author: KB). The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health. No competing financial interests exist.

References

- Beatty PC, & Willis GB (2007). Research Synthesis: the practice of cognitive interviewing. *Public Opinion Quarterly*, 71(2), 287–311. 10.1093/poq/nfm006.
- Benevides TW, Shore SM, Palmer K, Duncan P, Plank A, Andresen ML, Caplan R, Cook B, Gassner D, Hector BL, Morgan L, Nebeker L, Purkis Y, Rankowski B, Wittig K, & Coughlin SS (2020). Listening to the autistic voice: Mental health priorities to guide research and practice in autism from a stakeholder-driven project. *Autism*, 24(4), 822–833. 10.1177/1362361320908410. [PubMed: 32429818]
- Bujang MA, & Baharum N (2017). A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review. *Archives of Orofacial Science*, 12(1): 1–11.
- Bukavina L, Zaramo CEB, Tarabonata L, & Modlin CS (2017). Perceptions of Health-Related Quality of Life: a minority men's perspective. *American Journal of Men's Health*, 11(4), 990–998. 10.1177/1557988315580131.
- Diener E, Emmons RA, Larsen RJ, & Griffin S (1985). The satisfaction with Life Scale. *Journal of Personality Assessment*, 49(1), 71–75. 10.1207/s15327752jpa4901_13. [PubMed: 16367493]
- Diener (Ed.). (2022). (n.d.), Subjective well-being Retrieved June 3, from http://labs.psychology.illinois.edu/~ediener/Documents/Diener_1984.pdf
- Fang J, Fleck MP, Green A, McVilly K, Hao Y, Tan W, Fu R, & Power M (2011). The response scale for the intellectual disability module of the WHOQOL: 5-point or 3-point? *Journal of Intellectual Disability Research*, 55(6), 537–549. 10.1111/j.1365-2788.2011.01401.x. [PubMed: 21435065]
- Gliner JA, Morgan GA, & Harmon RJ (2001). Measurement reliability. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40(4), 486–488. 10.1097/00004583-200104000-00019. [PubMed: 11314576]
- Hong J, Bishop-Fitzpatrick L, Smith LE, Greenberg JS, & Mailick MR (2016). Factors Associated with subjective quality of life of adults with Autism Spectrum Disorder: self-report versus maternal reports. *Journal of Autism and Developmental Disorders*, 46(4), 1368–1378. 10.1007/s10803-015-2678-0. [PubMed: 26707626]
- Koo TK, & Li MY (2016). A Guideline of selecting and reporting Intraclass correlation coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155–163. 10.1016/j.jcm.2016.02.012. [PubMed: 27330520]
- Lam GYH, Sabnis S, Valcarlos MM, & Wolgemuth JR (2021). A critical review of academic literature constructing well-being in autistic adults. *Autism in Adulthood*. 10.1089/aut.2020.0053.
- McConachie H, Mason D, Parr JR, Garland D, Wilson C, & Rodgers J (2018). Enhancing the validity of a quality of life measure for autistic people. *Journal of Autism and Developmental Disorders*, 48(5), 1596–1611. 10.1007/s10803-017-3402-z. [PubMed: 29188584]
- McConachie H, Wilson C, Mason D, Garland D, Parr JR, Rattazzi A, Rodgers J, Skevington S, Uljarevic M, & Magiati I (2020). What is important in measuring quality of life? Reflections by autistic adults in four countries. *Autism in Adulthood*, 2(1), 4–12. 10.1089/aut.2019.0008. [PubMed: 36600984]

- Ne'eman A (2010). The future (and the past) of Autism Advocacy, or why the ASA's Magazine, the advocate, wouldn't publish this piece. *Disability Studies Quarterly*, 30(1), 10.18061/dsq.v30i1.1059.
- Nicolaidis C, Raymaker DM, McDonald KE, Lund EM, Leotti S, Kapp SK, Katz M, Beers LM, Kripke C, Maslak J, Hunter M, & Zhen KY (2020). Creating Accessible Survey Instruments for Use with autistic adults and people with intellectual disability: Lessons learned and recommendations. *Autism in Adulthood*, 2(1), 61–76. 10.1089/aut.2019.0074. [PubMed: 32355908]
- Pellicano E, & den Houting J (2022). Annual Research Review: shifting from 'normal science' to neurodiversity in autism science. *Journal of Child Psychology and Psychiatry*, 63(4), 381–396. 10.1111/jcpp.13534. [PubMed: 34730840]
- Power MJ, Green AM, & The WHOQOL-Dis Group. (2010). Development of the WHOQOL disabilities module. *Quality of Life Research*, 19(4), 571–584. 10.1007/s11136-010-9616-6. [PubMed: 20217246]
- PROMIS[®] Instrument Development and Validation Scientific Standards, Version 2.0, (revised May 2013). (n.d.). 72.
- Robertson SM (2010). Neurodiversity, Quality of Life, and autistic adults: shifting Research and Professional Focuses onto Real-Life Challenges. *Disability Studies Quarterly*, 30(1), 10.18061/dsq.v30i1.1069.
- Shoukri MM, Asyali MH, & Donner A (2004). Sample size requirements for the design of reliability study: review and new results. *Statistical Methods in Medical Research*, 13(4), 251–271. 10.1191/0962280204sm365ra.
- Skevington SM, Lotfy M, & O'Connell KA (2004). The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL Group. *Quality of Life Research*, 13(2), 299–310. 10.1023/B:QURE.0000018486.91360.00. [PubMed: 15085902]
- van Heijst BF, & Geurts HM (2015). Quality of life in autism across the lifespan: a meta-analysis. *Autism*, 19(2), 158–167. 10.1177/1362361313517053. [PubMed: 24443331]
- World Health Organization. Division of Mental Health (1996). WHOQOL-BREF: Introduction, administration, scoring and generic version of the assessment: field trial version, December 1996 (WHOQOL-BREF). World Health Organization. <https://apps.who.int/iris/handle/10665/63529>

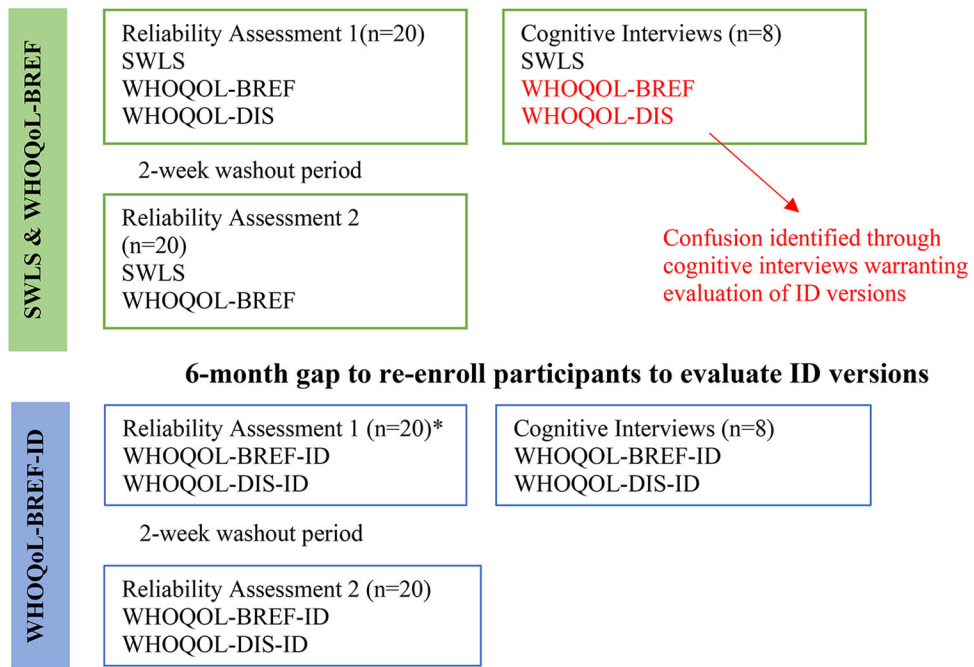


Fig. 1. Cohort prospective study design including reliability assessments and cognitive interviews per scale

Table 1

Reliability and understanding statistics for QOL and life satisfaction instruments

Measure	Mean (SD)	Internal Consistency (α)	Test-retest (ICC; 2,1)	Measure Understanding (% Agreement)
SWLS	27.8 (4.3)	0.789	0.948	100% (n = 8)
WHOQoL-BREF				
Physical Health Domain	16.25 (1.56)	0.387	0.89	
Without Q4		0.728	0.90	
Psychological Domain	15.7 (2.59)	0.851	0.94	
Social Relationships Domain	15.18 (2.86)	0.548	0.907	
Environmental Domain	15.8 (2.02)	0.791	0.83	
WHOQoL-BREF Total	62.8 (8.32)	0.911	0.95	75% (n = 6)
WHOQoL-DIS				
Discrimination Domain	11.6 (2.01)	0.352	0.839	
Autonomy Domain	12.3 (2.25)	0.836	0.768	
Inclusion Domain	22.65 (4.6)	0.853	0.870	
WHOQoL-DIS Total	46.5 (7.7)	0.881	0.882	75% (n = 6)
WHOQoL-BREF-ID (5-point)				
Physical Health Domain	16.03 (2.1)	0.464	0.877	
Without Q4		0.704	0.919	
Psychological Domain	16.2 (2.98)	0.821	0.93	
Social Relationships Domain	16.15 (3.1)	0.558	0.879	
Environmental Domain	15.72 (2.2)	0.696	0.869	
WHOQoL-BREF-ID Total	63.7 (8.67)	0.901	0.932	100% (n = 8)
WHOQoL-DIS-ID				
Discrimination Domain	11.95 (2.2)	0.412	0.854	
Autonomy Domain	12.75 (2.14)	0.689	0.786	
Inclusion Domain	23.3 (5.14)	0.924	0.942	
WHOQoL-DIS-ID Total	48 (8.3)	0.911	0.932	100% (n = 8)

Cognitive interview item-level challenges, classifications, and understanding agreement ratings

Table 2

Measure	Item Number	Item Wording	# endorsing confusion	Description of Confusion	Classification Type	Classification Timing	Measure Understanding(%agreement)
SWLS	N/A						100%
WHOQoL-BREF	3	“To what extent do you feel that physical pain prevents you from doing what you need to do?”	3 of 8	“to what extent” phrase	Lexical- phrasing of item	Understanding the item	75%
WHOQoL-BREF	4	“How much do you need any medical treatment to function in your daily life?”	3 of 8	Unsure of what constituted medical treatment	Concept-construct	Understanding the item	75%
WHOQoL-BREF	21	“Are you satisfied with your sex life, or your relationship with your partner? <i>For example, your husband/ wife, boyfriend/girlfriend.”</i>	5 of 8	Satisfaction concept was confusing if it was not applicable	Concept – construct	Choosing response	75%
WHOQoL-BREF	25	“How satisfied are you with your transport?”	5 of 8	“transport” word	Lexical-phrasing of item	Understanding the item	75%
WHOQoL-DIS	28	“Do you feel that some people treat you unfairly?”	1 of 8	Unsure of “some”	Concept–construct	Choosing response	75%
WHOQoL-DIS	30	“Do you worry about what might happen to you in the future? <i>For Example, thinking about not being able to look after yourself or being a burden to others in the future.</i> ”	1 of 8	Meaning of “burden”; did not prefer the example	Lexical-phrasing of item example	Understanding the item	75%
WHOQoL-BREF-ID	21	“Are you satisfied with your sex life, or your relationship with your partner?”	5 of 8	Satisfaction concept was confusing if it was not applicable	Concept – construct	Choosing response	100%
WHOQoL-DIS-ID	N/A						100%