A systematic review of quality of life measures for people with intellectual disabilities and challenging behaviours

C. Townsend-White,1 A. N. T. Pham1 & M. V. Vassos1,2

1 Centre of Excellence for Behaviour Support, University of Queensland, Ipswich, Queensland, Australia
2 Discipline of Disability Studies, RMIT University, Bundoora, Victoria, Australia

Abstract

Background The quality of life (QOL) construct is proposed as a method to assess service outcomes for people utilising disability services. With this in mind, the aim of this study was to conduct a systematic review of available QOL measures for people with intellectual disability (ID) to pinpoint psychometrically sound measures that can be routinely used for service evaluation.

Method A systematic search of the disability literature published between 1980 and 2008 was conducted in order to identify appropriate QOL tools for use within an Australian context. Twenty-four QOL instruments were identified and each instrument was then evaluated against a set of psychometric and measurement criteria.

Results Six of the instruments examined were deemed to be psychometrically sound on the available information. No instruments were found that specifically assess QOL for people with ID who exhibit challenging behaviour. Most of the instruments assess QOL from a subjective perspective, use a questionnaire format and measure only some (not all) of the eight theoretically accepted domains of QOL.

Conclusions More instruments that measure QOL need to be developed and rigorously validated. This is especially the case for high-needs disability populations like those individuals that exhibit challenging behaviour or have severe to profound ID, as it is questionable whether existing measures can be used with these populations.

Keywords challenging behaviour, intellectual disabilities, measurement, quality of life

Background The Convention on the Rights of Persons with Disabilities (United Nations 2006) acknowledges that people with disabilities continue to face rights violations and barriers to equal participation in society. There is growing recognition of the need to ensure that all people who have disabilities enjoy human
rights on an equal basis with others. These rights pertain to a range of life domains. While there is no international consensus on what constitutes a human rights approach to social policy and services (Gruskin & Loff 2002), human rights treaties and documents explicitly or implicitly include the right of people with a disability to expect and receive effective evidence-based services and care.

Regardless of a country’s human rights perspective or economic status, the lives of adults with ID, especially those who have severe and profound ID, mental illness and those who exhibit challenging behaviours, differ significantly from others in the community (Cummins 2005a; United Nations 2006). Historically, service systems have failed to adequately meet the needs of people with ID and have not met the obligations articulated within the Convention on the Rights of Persons with Disabilities. Services have been crisis-driven, ad hoc, have denied human dignity and have failed to facilitate or encourage meaningful community engagement. Service responses have included restrictive practices including: physical/mechanical and chemical restraint, and systematic neglect resulting in adverse physical and psychological effects for consumers and carers (Bernal & Hollins 1995; Deveau & McGill 2009; Sanders 2009).

Within a human rights framework people with intellectual disability (ID) have the right to accessible, appropriate, evidence-based services that enable them to achieve personal goals and to enjoy a quality of life (QOL) equal to that of people who do not have a disability. There is a pressing need within the ID service system for more consistent and responsive services which facilitate positive outcomes. Quality-based reform movements within health, mental health, education and disability fields increasingly emphasise the importance of accountability and the establishment of quality frameworks that include reporting mechanisms to monitor service and resource effectiveness and congruence with consumer needs (Pirkis et al. 2005; Brown et al. 2009).

Quality monitoring systems are required in ID services as for any other disability group. Routine outcome measurement systems constitute an important part of a quality system and can promote effective evidence-based policy and service development, planning and delivery. Comprehensive outcome data include individual and population outcomes, service level outcomes including levels of service utilisation, efficiency, effectiveness, accessibility, equity and appropriateness (Bickman et al. 1998; Hensel 2001), and economic data including resource use and costs at the level of individual users, their families and the community.

A comprehensive outcome measurement system includes measures that assess individual consumer outcomes or changes in an individual, which are wholly attributable to the ID service or intervention. Human rights frameworks recognise the right of people with a disability to have full and effective participation equal to other members of society generally and particularly in relation to matters directly concerning them. Consumer outcome measurement must therefore include the perspective of the individual service user as well as family members and the service provider. Consumer outcome measurement should be ‘embedded’ in routine service delivery and data should be collected on entry to a service and at regular intervals while the individual is using the service. Consumer outcomes should be directly relevant to the goals articulated in individual service users’ plans including person-centred plans or positive behaviour support plans.

A key individual outcome domain is QOL. The concept of QOL has evolved within ID since the 1980s (Verdugo et al. 2005; Zuna et al. 2009). Areas of focus have shifted from the effects of institutionalisation and untenable environments to ‘higher order’ aspects of QOL (Cummins & Lau 2003). QOL is associated with human values including happiness, satisfaction, general feelings of well-being and opportunities to achieve personal potential (Cummins 1991; Schalock 2000; Brown & Brown 2009).

Quality of life is a multidimensional phenomenon composed of core domains and elements that are influenced by personal characteristics and environmental and contextual variables (Verdugo et al. 2005). Domains and specific indicators of QOL have been identified and critically assessed (Brown & Brown 2009). QOL domains include those that are common to all humans and additional domains that may be unique to the individual (Lyons 2005; Verdugo et al. 2005). There is international consensus about the dimensions of QOL (Schalock et al.)
2002; Brown & Brown 2009). Eight core QOL domains have been identified and validated in a series of cross-cultural studies: emotional well-being, interpersonal relationships, material well-being, personal development, physical well-being, self-determination, social inclusion and rights (Schalock 2000; Beadle-Brown et al. 2009; Wang et al. 2010). When aggregated, these domains represent the total QOL construct. QOL outcomes may be impacted upon by variables within several domains, for example, the nature of service delivery, the consumer’s life experience, their physical environment, and the number and nature of people with whom the consumer lives and with whom they interact. Individual QOL varies over time and is substantially influenced by cultural environments, individual interests and circumstances. As a measurement, QOL should be considered holistically and family QOL should also be considered (Schalock et al. 2009).

Some aspects of QOL can be objectively measured through direct observation and assessed and verified against commonly held standards, for example, medical health and wealth (McGillivray et al. 2009). Other aspects of QOL can only be assessed from the subjectively perceived experience of the individual through self-report. QOL domains should therefore include separately measured subjective and objective indicators (Verdugo et al. 2005; Beadle-Brown et al. 2009). It is important to note that subjective and objective indicators are typically not highly correlated (Schalock et al. 2002, 2008; Cummins 2005b).

As a consumer outcome measure, QOL enables impact studies, cross-cultural and country comparisons, and provides a framework for service principles and service design, delivery and evaluation that is congruent with human rights frameworks (Schalock 2004; Schalock et al. 2008; Brown et al. 2009). QOL outcome data can inform a range of stakeholders as to whether and to what extent service users experience changes in their QOL as a result of service interventions and whether personal outcomes are being achieved (Bonham et al. 2004; Brown & Brown 2005; Schalock et al. 2008). It can guide quality improvement strategies at individual, service and community levels (Schalock et al. 2002, 2008; Hoffman et al. 2006; D’Eath et al. 2009; Kober & Egginton 2009) and influence political development and reform (Cummins & Lau 2003).

Quality of life is developing from a concept to a measurable and socially valid construct (Claes et al. 2009). There is growing acceptance of the need for QOL to be measured, over time and in relation to services received (Brown et al. 2009; Claes et al. 2009; Griffith & Owen 2009; Isaacs et al. 2009; Murphy 2009). Hughes et al. (1995; cited in Claes et al. 2009) found over 1243 measures of QOL reported in the ID literature by the mid-1990s; however, there are no systematic reviews or comparative analyses of QOL measures for people with ID against key criteria.

Aim

The aim of this paper is to identify and systematically review QOL measures that could be used routinely by researchers and service providers in measuring subjective and objective QOL for people who have an ID. Instruments that specifically measured the QOL of people with ID and challenging behaviours were also sought in this review process.

Methodology

The approach taken in this review process involved recourse to the published scientific literature, the ‘grey’ literature and consultation with experts in the field. Searches were limited to publication dates from 1980 to 2008. This time is consistent with the consensus within the literature that research into QOL for vulnerable groups commenced in the 1980s (Verdugo et al. 2005; Schalock et al. 2008; Zuna et al. 2009). Existing systematic reviews of QOL instruments were initially sought to identify appropriate ID keywords. Appropriate search terms were identified by exploring the MeSH dictionary and thesaurus definitions in the Cochrane library database. No systematic reviews relating to QOL measurement tools for people with ID were located. Cochrane reviews by Balogh et al. (2008) and Has-siotis & Hall (2008) relating to ID were obtained to identify key terms and descriptors in the current literature.

After consultation with qualified archivists, search strings, specific to each database (including key-
words, MeSH terms, EMTREE terms, explode terms and proximity operators), were devised and validated against a synonym ring of terms (Sanderson 2006) designed to extract published material relating to people with ID in existing databases (see Appendix 1 for these search strings and terms). A search of the peer-reviewed literature was then undertaken to identify studies that examined QOL in individuals with ID and challenging behaviours. Figure 1 depicts the literature search process that was utilised.

A systematic search was conducted of Medline, PsycINFO, ERIC and CINAHL electronic databases. The initial search identified 3081 papers. The identified articles from each database were imported into Endnote X® databases and duplicates were deleted, reducing the number of papers to 2951. Article titles were then scanned for relevance, further reducing the number of papers to 404. In the initial screening phase, titles that mentioned QOL measurement and indicated that the study sample was composed of individuals with ID were retained. Abstract summaries were also read in the screening phase to provide background information and objectives of the study, eligibility criteria, data sources and participant information. The 404 relevant articles were read in full by at least one author. Articles that did not focus on domain measurement of QOL in people with learning, developmental or ID were culled. This reduced the total number of articles to 273. Experts were contacted for further data sources and the grey literature was searched for additional reports and papers. Scholarly databases, resource catalogues and websites of government agencies and non-government organisations were also searched to identify reports on QOL outcome measures. Additional data were identified through reference lists of review articles, recommended articles from experts and other electronic sources. These articles were hand searched to supplement any studies that may not have been identified by the database search. No additional useful papers were identified.

The remaining 273 articles were read in full. The articles were further reduced based on the following exclusion criteria: purely health-related QOL measures, conceptual papers and books, book sections without data on the QOL measure/s, observational studies, behavioural matrixes, functional behavioural assessments, checklists and inventories, diagnostic instruments, instruments developed for a specific age range or restrictive sample. This excluded 142 articles from the review. Using this process, 24 QOL instruments were identified. Table 1 lists these measures.

**Instrument extraction**

The 24 instruments were examined and excluded from the review if they demonstrated the following characteristics: (1) does not assess domains that align with the eight core domains of QOL (Schalock & Verdugo 2005); (2) take longer than 2 h to administer; (3) no psychometric data available; and (4) instruments in languages other than English.

**Evaluating the measures**

A search of ID and mental health literature was undertaken and a set of criteria were developed against which measures should be assessed for inclusion as potential routine measures. These criteria included: (1) the measurement and aggregation of objective and subjective dimensions across a range of life domains should be used to ascertain a total QOL construct; (2) instruments should have good psychometric properties, be brief, simple to administer and easy to score; (3) a pre-test should be used to establish that the respondents can comprehend the questions; and (4) instruments should have a consumer perspective, be acceptable to consumers and promote dialogue between consumers and providers (Andrews et al. 1994; Bickman et al. 1997; Bickman et al. 1998; Cummins 2001; Claes et al. 2009). Based on this review, we developed 24 assessment criteria that were used to review each of the relevant instruments. Criteria are described in more details below.

**Reliability**

Internal consistency coefficients (Cronbach’s α) were examined to determine whether scale items were measuring the same domain. Traditional acceptable alpha coefficients were used in the review; however, instruments with reported alphas approaching 0.7 were also considered. The reviewers considered test–retest reliability to assess if the
Figure 1: Flowchart of the systematic literature search protocol. QOL, quality of life; ID, intellectual disability; LD, learning disability; LOTE, languages other than English; HRQOL, health-related quality of life.

© 2011 The Authors. Journal of Intellectual Disability Research © 2011 Blackwell Publishing Ltd
instrument produced consistent results for the same participant on different administrations. No consistent test–retest timeframe was used to exclude papers based on this criterion as testing timeframes varied significantly in the literature collated. Kappa coefficients and intra-class correlation coefficients were reviewed to determine if the measure reliably assessed QOL ratings between raters. Acceptable criteria for inter-rater agreement were based on kappa coefficients criteria (>0.6) reported by Streiner & Norman (2003). Intra-class correlation coefficients were also accepted as a test of inter-rater reliability, with correlations above 0.60 deemed to be acceptable (Ouellette-Kuntz 1990).

Validity

Criterion validity is determined by assessing a new measure against a previously validated measure. This concurrent form of validity is demonstrated by comparing the new instrument with the gold standard of assessment in the field. Rapley et al. (1997) noted that many validation studies assessed their developed measures against the Schalock Quality of Life Questionnaire (QOL-Q; Schalock & Keith 1993). Criterion validity is demonstrated if the domains in the new measure correlated highly with the same domains on a previously validated measure of QOL.

Construct validity was assessed in three ways; one was to examine factor-analytical studies to determine whether any studies disconfirmed the theorised structure of the measure. The second method looked at convergent validity results, that is, the degree to which an item is related to the criterion that theoretically represents the same or similar construct. The third method looked at discriminant validity results, and assessed the measure against a criterion measure with which it should not be associated. In the review, correlations between theoretically linked scales and sub-scales were examined. Moderate to high correlations indicated acceptable convergent validity. Discriminant validity was

Table 1 List of the 24 measures considered for evaluation

<table>
<thead>
<tr>
<th>Measure</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choice Questionnaire</td>
<td>Standcliffe &amp; Parmenter (1999)</td>
</tr>
<tr>
<td>Comprehensive Quality of Life Scale (Com-QOL-ID)</td>
<td>Cummins (1991)</td>
</tr>
<tr>
<td>Evaluation of Quality of Life Instrument (EQLI)</td>
<td>Nata et al. (2006)</td>
</tr>
<tr>
<td>Life Satisfaction Matrix</td>
<td>Lyons (2005)</td>
</tr>
<tr>
<td>Lifestyle Satisfaction Scale (LSS)</td>
<td>Heal &amp; Chadsey-Rusch (1985)</td>
</tr>
<tr>
<td>Quality of Life Questionnaire</td>
<td>Brown &amp; Bayer (1992)</td>
</tr>
<tr>
<td>Multifaceted Life Satisfaction Scale (MLSS)</td>
<td>Harner &amp; Heal (1993)</td>
</tr>
<tr>
<td>Quality of Life, Enjoyment and Satisfaction Questionnaire (Q-LES-Q)</td>
<td>Endicott et al. (1993)</td>
</tr>
<tr>
<td>Quality of Life Interview Schedule (QUOLIS)</td>
<td>Ouellette-Kuntz (1990)</td>
</tr>
<tr>
<td>Quality of Life Questionnaire (QOL-Q)</td>
<td>Schalock &amp; Keith (1993)</td>
</tr>
<tr>
<td>Questionnaire on Quality of Life</td>
<td>Cragg &amp; Harrison (1986)</td>
</tr>
<tr>
<td>Guernsey Community Participation and Leisure Assessment (GCPLA)</td>
<td>Baker (2000)</td>
</tr>
<tr>
<td>Maryland Ask Me! Project</td>
<td>Bonham et al. (2004)</td>
</tr>
<tr>
<td>Personal Outcomes Measure</td>
<td>Gardner et al. (1997)</td>
</tr>
<tr>
<td>Overt Behaviour Scale</td>
<td>Kelly et al. (2006)</td>
</tr>
<tr>
<td>Life Experiences Checklist</td>
<td>Ager &amp; Eglington (1989)</td>
</tr>
<tr>
<td>Quality of Life Index</td>
<td>Campo et al. (1996)</td>
</tr>
<tr>
<td>Quality of Life Assessment Tool</td>
<td>Johnson &amp; Cocks (1989)</td>
</tr>
<tr>
<td>Quality of Life Instrument</td>
<td>Janssen et al. (1999)</td>
</tr>
<tr>
<td>Consumer Satisfaction Survey</td>
<td>Temple University (1988)</td>
</tr>
<tr>
<td>Quality of Life Assessment Form</td>
<td>McGuire et al. (1991)</td>
</tr>
<tr>
<td>Quality of Life Instrument Package</td>
<td>Raphael et al. (1999)</td>
</tr>
<tr>
<td>Mood Interest Pleasure Questionnaire</td>
<td>Ross &amp; Oliver (2003)</td>
</tr>
</tbody>
</table>

© 2011 The Authors. Journal of Intellectual Disability Research © 2011 Blackwell Publishing Ltd
assessed by checking if unrelated items produced low correlations.

Content validity was evaluated to determine whether the instrument items had been generated in accordance with relevant theory. The item development process was reviewed to determine whether the instrument content covered a representative sample of the domain that was being measured. High face validity has been found to increase participation rates, and increase professional care staff and clinicians’ perceived relevance of the measure, induce cooperation in participants and promote acceptance of measures by policy makers and bureaucrats (Nevo 1985; McLeod 2003). QOL measures were read to check for face validity, that is, whether an instrument appeared to measure what it claimed to measure and whether it was acceptable to respondents.

Acquiescence

Response bias has been well documented in self-report measures among people with ID (Siegelman et al. 1980; Verri et al. 1999). Adolescents and adults with ID are susceptible to presentation bias and have a tendency to choose the most positive response alternative (Verri et al. 1999; Schalock et al. 2002). The authors examined whether the instrument included screening tools, pre-tests or acquiescence scales to assess a person’s capacity to understand the questions.

Norms and pilot testing

Norm data were collected to determine if the instrument had been piloted using an Australian sample. Cultural biases were assessed through cross-cultural studies where available. Published norms for a representative population and the target population were sought for comparison. Information was sought on how the scale was developed with a focus on whether prior testing had been conducted to determine if the instrument would be suitable for people with ID.

Administration

Measures were also considered in terms of their affordability. Existing QOL measures were examined for the cost of an introductory kit/manual and price per 10 forms in Australian dollars. Instrument brevity is an important feature of outcome measures. Briefer instruments reduce data collection burden thereby promoting routine collection by clinicians (Bickman et al. 1998; Hermann et al. 2000). Response formats have a significant effect on the quality of the data collected. Hartley & MacLean (2006) conducted a review of the literature on the reliability and validity of Likert-type scales for people with ID and found that the level of ID significantly affects the reliability of responses on Likert-type scales. Individuals with borderline to mild ID responded more consistently to Likert scales than those with moderate to severe ID. Data were collected on any training or specialist knowledge required to determine ease of administration and applicability in QOL evaluation. Measures were also evaluated for reading level (grade required to complete the test) where the data were available.

Results

No specific instruments that measure the QOL of people with ID and challenging behaviour were found. Six instruments were assessed psychometrically in terms of the 24 selection criteria. Tables 2 and 3 provide a summary of these instruments listed in descending order in terms of their strength against the criteria. The Multifaceted Lifestyle Satisfaction Scale (MLSS; Harner & Heal 1993) was deemed the most psychometrically sound instrument with evidence of adequate internal consistency, test–retest reliability, inter-rater reliability and construct validity. While the MLSS shows promise, the measure has only been validated by the authors; further validation studies are required to replicate the initial findings. Lack of replication is not just present for the MLSS, lack of replication and generally, a lack of available information plagued the instrument review process.

One limitation of the MLSS (Harner & Heal 1993) is that it does not measure all of the theoretically accepted eight domains of QOL. Only one measure, the QOL-Q (Schalock & Keith 1993), is designed to assess these eight domains; however, the Social Belonging/Community Integration subscale of the instrument does not demonstrate sufficient reliability in validation studies. Another
Table 2 Quality of life (QOL) instruments assessed against 24 review criteria

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Y – domains include:</td>
<td>1. community satisfaction</td>
<td>1. material well-being</td>
<td>1. overall satisfaction</td>
<td>1. standard of living</td>
<td>Y – domains include:</td>
<td>Y – domains include:</td>
</tr>
<tr>
<td>N – domains that do not align with the eight core domains of QOL</td>
<td>2. friends and free time</td>
<td>2. health</td>
<td>2. health</td>
<td>2. family and guardianship</td>
<td>with the eight core domains of QOL</td>
<td>1. quality of service received</td>
</tr>
<tr>
<td>N – assesses domains that do not align with the eight core domains of QOL</td>
<td>3. satisfaction with interpersonal interactions</td>
<td>3. productivity</td>
<td>3. life achievement</td>
<td>3. income maintenance</td>
<td></td>
<td>2. satisfaction with opportunities for social interaction</td>
</tr>
<tr>
<td>N – assesses domains that do not align with the eight core domains of QOL</td>
<td>4. recreation and leisure</td>
<td>4. intimacy</td>
<td>4. personal relationships</td>
<td>4. education, training and employment</td>
<td></td>
<td>3. satisfaction with living environment</td>
</tr>
<tr>
<td>N – assesses domains that do not align with the eight core domains of QOL</td>
<td>5. satisfaction with services</td>
<td>5. safety</td>
<td>5. personal safety</td>
<td>5. housing and safety</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N – assesses domains that do not align with the eight core domains of QOL</td>
<td>6. place in community</td>
<td>6. social belonging and community integration</td>
<td>6. feeling part of the community</td>
<td>6. transportation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N – assesses domains that do not align with the eight core domains of QOL</td>
<td>7. emotional well-being</td>
<td>7. future security</td>
<td>7. case management</td>
<td>7. social and recreational</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N – assesses domains that do not align with the eight core domains of QOL</td>
<td>8. enjoyment of leisure</td>
<td>N/A</td>
<td>8. religious and cultural</td>
<td>8. advocacy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N – assesses domains that do not align with the eight core domains of QOL</td>
<td>9. case management</td>
<td>N/A</td>
<td>9. case management</td>
<td>9. social and recreational</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N – assesses domains that do not align with the eight core domains of QOL</td>
<td>10. advocacy</td>
<td>N/A</td>
<td>10. advocacy</td>
<td>10. social and recreational</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N – assesses domains that do not align with the eight core domains of QOL</td>
<td>11. counselling</td>
<td>N/A</td>
<td>11. counselling</td>
<td>11. social and recreational</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N – assesses domains that do not align with the eight core domains of QOL</td>
<td>12. aesthetics</td>
<td>N/A</td>
<td>12. aesthetics</td>
<td>12. social and recreational</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Informant

SR – self-report
P – proxy

Internal consistency

* Low (α < 0.60)
** Moderate (0.6 ≤ α ≤ 0.79)
*** High (α ≥ 0.8)

Test–retest reliability

T2 – tested within 2 weeks
Tm – multiple testing periods
T1 – other

Test–retest reliability coefficient

* Low (r < 0.70)
** Moderate (0.70 ≤ r ≤ 0.79)
*** High (r > 0.8)

Inter-rater reliability

* Poor (kappa < 0.40)
** Fair (0.40 ≤ kappa < 0.59)
*** Moderate (0.60 ≤ kappa ≤ 0.74; intra-class correlation > 0.60)
**** Very good (kappa ≥ 0.74)

Inter-rater reliability (cross informant agreement)

Y – yes
N – no

Content validity

Y – demonstrated test-item analysis, expert review or field trials with group or subgroup samples
N – items were not derived from theory

Construct validity

Y – factor-analytical studies conducted
N – no factor-analytical studies data available

© 2011 The Authors. Journal of Intellectual Disability Research © 2011 Blackwell Publishing Ltd
<table>
<thead>
<tr>
<th>Quality of Life Measure</th>
<th>Face Validity</th>
<th>Convergent Validity</th>
<th>Discriminant Validity</th>
<th>Acquiescence Scale</th>
<th>Norm Data</th>
<th>Cost</th>
<th>Administration Time (minutes)</th>
<th>Response Format</th>
<th>Administrator Training</th>
<th>Administration Method</th>
<th>Reading Level (Grade Level Required)</th>
<th>Level of Disability Appropriate for Administration</th>
<th>QOL Type</th>
<th>Pilot Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multifaceted Life Satisfaction Scale (MLSS; Harner &amp; Heath 1993)</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>20</td>
<td>GAS [58 items, +2 (enthusiastic yes) to -2 (enthusiastic no)]</td>
<td>N/A</td>
<td>N/A</td>
<td>S</td>
<td>N/A</td>
<td>Mild/moderate/severe/profound</td>
<td>SWB</td>
</tr>
<tr>
<td>Comprehensive Quality of Life Scale (Com-QOL-ID; Cummins 1997)</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>N/A</td>
<td>Y</td>
<td>$110</td>
<td>20</td>
<td>Lk (Q1 items, rated on multiple response formats)</td>
<td>N/A</td>
<td>N/A</td>
<td>S</td>
<td>S</td>
<td>Mild</td>
<td>OWB/SWB</td>
</tr>
<tr>
<td>Quality of Life Questionnaire (QOL-Q; Schalock &amp; Keith 1993)</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>N/A</td>
<td>Y</td>
<td>Free – with author’s permission</td>
<td>20–30</td>
<td>Lk (40 items, 1 (low) to 3 (high))</td>
<td>45</td>
<td>N/A</td>
<td>S</td>
<td>S</td>
<td>N/A</td>
<td>SWB/OWB</td>
</tr>
<tr>
<td>Personal/Wellbeing Index (PWI-ID; Cummins &amp; Lau 2005)</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Free – with author’s permission</td>
<td>N/A</td>
<td>Lk (seven items, rated on multiple response formats)</td>
<td>N/A</td>
<td>N/A</td>
<td>S</td>
<td>S</td>
<td>N/A</td>
<td>SWB</td>
</tr>
<tr>
<td>Quality Of Life Interview Schedule (QUOLIS; Ouellette-Kuntz 1990)</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>$110</td>
<td>Free – with author’s permission</td>
<td>Lk (7 items, 1 (terrible) to 7 (delighted))</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Mild/moderate/severe</td>
<td>OWB</td>
</tr>
<tr>
<td>Evaluation of Quality of Life Instrument (EQLI; Nota et al. 2006)</td>
<td>N/A</td>
<td>Y</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Free – with author’s permission</td>
<td>N/A</td>
<td>Lk (18 items, 1 (does not describe him or her at all) to 5 (describes him or her perfectly well))</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>OWB</td>
</tr>
</tbody>
</table>

N/A refers to information being not available.
The limitation of the measure is the combination of objective and subjective items causing the scale to correlate more strongly with objective rather than subjective QOL variables (Cummins 1997).

Only one instrument, the Comprehensive Quality of Life Scale (Com-QOL-ID; Cummins 1997) attempted to measure QOL domains on objective and subjective QOL on separate axes as recommended in the literature; however, the use of this scale cannot be recommended as the 21 items of the objective scale did not factor as intended in validation studies and the use of importance as a weighting for client satisfaction is invalid (Cummins 1997). As a result of the limitations, the authors have abandoned the measure and installed in its place the Personal Wellbeing Index (PWI-ID). The PWI-ID has similar psychometric properties as the Com-QOL-ID as the items are derived from the Com-QOL-ID; however, the PWI-ID in its current form has not been empirically validated.

C. Townsend-White et al. • A systematic review of quality of life measures

Table 3 Extended summary of the six quality of life instruments identified in terms of reliability and validity information available

<table>
<thead>
<tr>
<th>Measure</th>
<th>Reliability</th>
<th>Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multifaceted Life Satisfaction Scale (MLSS)</td>
<td>Test–retest correlations (multiple time periods utilised and two participant samples used) for the domains were mainly above 0.60 with the exception of general satisfaction and control and determination. Cronbach ( \alpha ) statistics across the domain were also mainly above 0.70 with the exception of general satisfaction and control and determination and friends and free time. Inter-rater reliability for the overall MLSS scale was 0.99 and ranged from 0.81 to 0.97 for the domains. Convergent validity was exhibited by significant correlations between Personal Caretaker Questionnaire domain scores and MLSS domain scores. Fewer significant correlations were found between the MLSS domain scores and Quality of Life Questionnaire (Schalock et al. 1990, cited in Harner &amp; Heal 1993) domain scores.</td>
<td></td>
</tr>
<tr>
<td>Comprehensive Quality of Life Scale (Com-QOL-ID)</td>
<td>Cronbach ( \alpha ) statistics ranging from 0.22 to 0.65 for each of the objective axis domain; ( \alpha ) for the total objective scale was 0.56. Regarding the subjective axis, Cronbach ( \alpha ) was 0.68 for the satisfaction sub-scale and 0.48 for the importance sub-scale. Test–retest reliabilities (1–2 weeks) for the seven domains in the importance scale ranged between 0.12 and 0.97 for the importance scale and 0.13 and 0.86 for the satisfaction scale. Consistent with QOL literature, no significant correlation between objective and subjective axis scores were found. The exception to this was objective health which correlated positively with health importance and health satisfaction. These results highlight construct validity.</td>
<td></td>
</tr>
<tr>
<td>Quality of Life Questionnaire (QOL-Q)</td>
<td>Cronbach ( \alpha ), alpha for the domains ranged from 0.60 and 0.82 with ( \alpha = 0.90 ) for the total score (Schalock &amp; Keith 1993). Inter-rater reliabilities for the domains ranged from 0.67 to 0.90 (Schalock &amp; Keith 1993). The QOL-Q has been tested in an Australian sample and found to be psychometrically sound (Kober &amp; Eggleton 2009). The QOL-Q has been used as a basis for creating a cross-cultural QOL measure (Schalock et al. 2005; Chou et al. 2007); however, limited psychometric data are available for this measure (some face validity results and divergent validity).</td>
<td></td>
</tr>
<tr>
<td>Personal Wellbeing Index (PWI-ID)</td>
<td>Cronbach ( \alpha ) has been reported as 0.76, with a 1- to 2-week test–retest reliability of 0.58 at 1–2 weeks (McGillivray et al. 2009). Australian norms are available for the PWI-ID. Construct validity was assessed by McGillivray et al. (2009) who confirmed the factor structure of the PWI-ID using principal components analysis.</td>
<td></td>
</tr>
<tr>
<td>Quality of Life Interview Schedule (QUOLIS)</td>
<td>65% of the QUOLIS scores have acceptable inter-rater agreement (intra-class correlations above 0.60) across three independent raters. Inconsistent inter-rater agreement appeared in the scoring of the family/guardianship, income maintenance and counseling domains. No validity information is available.</td>
<td></td>
</tr>
<tr>
<td>Evaluation of Quality of Life Instrument (EQLI)</td>
<td>Cronbach ( \alpha ) for the domains range from 0.73 to 0.85 (Nota et al. 2006). Evidence for the construct validity was obtained through exploratory factor analyses which confirmed the three dimensions of the EQLI. Convergent validity was demonstrated via significant correlations above 0.30 between EQLI scores and Quality of Life Index scores (Keith et al. 1986, cited in Nota et al. 2006). EQLI scores were also found to discriminate between groups on the basis of social abilities and psycho-pathological problems, hence demonstrating adequate divergent validity.</td>
<td></td>
</tr>
</tbody>
</table>

© 2011 The Authors. Journal of Intellectual Disability Research © 2011 Blackwell Publishing Ltd
Two of the reviewed measures relied predominantly on proxy responses. The Quality of Life Interview Schedule (QUOLIS; Ouellette-Kuntz 1990) is a structured interview designed to measure the QOL of adults who are unable to complete a written or verbal interview, that is, adults without verbal communication skills or have severe ID. Thus, data are provided by proxy respondents, although the measure contains a section the ‘Residential Satisfaction Questionnaire’ that can be answered by verbal clients. The Evaluation of Quality of Life Instrument (EQLI; Nota et al. 2006) is designed to elicit staff evaluations of the level of satisfaction experienced by the individuals they support. Validation studies suggest that the measure is an effective tool for gaining staff perceptions of individuals who may be at risk of dissatisfaction with various QOL domains.

Three of the six measures have the capacity to check if the individual being examined has a tendency for acquiescent responding. Acquiescent responding (i.e. socially biased responses) is something that all researchers and service providers working with people with ID need to be aware of when conducting research or service reviews. The MLSS (Harner & Heal 1993) assessed acquiescence which is another advantage of this measure when compared to the others; however, the method of assessing acquiescent responding is not as rigorous as the method utilised by the Com-QOL-ID (Cummins et al. 1997) and the PWI-ID (Cummins & Lau 2005). These measures use a structure pre-test procedure which the individual must complete successfully before they can proceed to completing the measure whereas the MLSS has an inbuilt acquiescent responding scale comprised of existing items in the scale. There is no evidence that one form of assessing acquiescence is better than the other on the basis of the literature reviewed for this study.

Discussion

Validly measured subjective and objective QOL is an important aspect of consumer outcome measurement in the area of ID. QOL instruments can inform governments and service providers whether services are meeting the needs of the population they are intended to serve and to assist individuals and their families and carers to ascertain whether their needs are being met to their satisfaction over time (Bickman et al. 1998). Reliable, routinely collected consumer QOL data can inform policy and service development and reform. The findings of this study establish that there are several QOL measures for people with ID. Of these, a limited number are psychometrically sound and potentially suitable for routine measurement of QOL among people with ID according to criteria we developed from consumer outcome measurement literature. However, no universally accepted ‘gold standard’ instrument exists for assessing the QOL of adults with ID who exhibit challenging behaviour.

Routine outcome measurement is increasingly being employed to assess service effectiveness in a range of sectors, including health and mental health. Personal outcomes can be analysed at the level of the individual, aggregated at the organisation or systems level and complemented by other performance measures such as health and safety indicators, client movement patterns, staff turnover and unit costs (Gardner & Carran 2009). The utility of the QOL concept is still emerging in the field of ID and there is modest literature on its application and extent of impact (Brown & Brown 2005; Brown et al. 2009). Research in the area of QOL has moved from conceptualisation to a second phase of research that includes understanding of the emic and etic properties of QOL: homeostatic theory of subjective well-being (Cummins 2002) and the promotion of QOL through the identification and articulation of QOL aspects of policy and consequent integration of QOL measurement in policy and planning (Verdugo & Schalock 2009).

Continued research efforts are needed to develop ‘one or more detailed and empirically based QOL models and/or theories that include inputs, outputs and a clear delineation of mediator and moderator variables’ (Verdugo & Schalock 2009, p. 63), including those appropriate to people with ID and challenging behaviour. An increased number and variety of instruments will enable researchers and clinicians to apply appropriate instruments based on the aims of their services and projects and the resources at their disposal.

Governments need to fund research and demonstration projects to develop culturally sensitive QOL
models and data sets that monitor and evaluate the effects and impact of QOL-orientated policies and practices. Researchers and evaluators in turn need to foster the use of a holistic approach to QOL research by incorporating participatory action research and evaluation to ensure that QOL concepts reflect the subjective experience of people with ID including those who exhibit challenging behaviour and facilitate their participation in evaluation (Bonham et al. 2004).

A cautionary word should perhaps conclude. While QOL is an important aspect of outcome measurement, it should not be the sole focus of outcome measurement systems. The ‘tyranny of quality of life’ should be avoided and QOL should be one of a raft of measures that make up a comprehensive measurement system (Kober & Eggleton 2009, p. 47). The dangers in overly focusing on QOL include: failure to acknowledge the subjective experience of people with ID, over reliance on a single measure and failure to recognise domains or elements of significant deprivation within composite scoring (Isaacs et al. 2009). In addition, understanding QOL levels for people with ID must be considered in relation to the wider population. While scores may increase over time for people with ID, they may still remain below a level acceptable to other groups in the community including people with other disabilities and those who do not have a disability.

Acknowledgements, source of funding and conflict of interest
The authors would like to thank Prof. Leonard Bickman for his valuable comments on a previous draft of this research paper. No external funding was utilised to undertake this research. No conflicts of interest were present during this study.

References


A systematic review of quality of life measures


Appendix I

Search terms utilised across the four literature databases searched

<table>
<thead>
<tr>
<th>Database</th>
<th>Search terms</th>
</tr>
</thead>
</table>
| Medline (via Ovid) | 1. mentally disabled persons.mp.  
2. mental retardation.mp.  
3. disabled persons.mp.  
4. treatment outcome$.mp.  
5. quality of life.mp.  
6. consumer$. outcome$.mp.  
7. client outcome$.mp.  
8. client satisfaction.mp.  
9. outcome$.mp.  
10. outcome instrument$.mp.  
11. assess$.mp.  
12. scale$.mp.  
13. checklist$.mp.  
14. inventor$.mp.  
15. rating$.mp.  
16. schedule$.mp.  
17. happiness Ind$.mp.  
18. instrument$.mp.  
19. subjective wellbeing.mp.  
20. objective wellbeing.mp.  
21. emotional wellbeing.mp.  
22. physical wellbeing.mp.  
23. material wellbeing.mp.  
24. domain$.mp.  
25. social inclusion.mp.  
26. community participation.mp.  
27. self determination.mp.  
28. self determination.mp.  
29. option$.mp.  
30. rights.mp.  
31. safe.mp.  
32. objective.mp.  
33. subjective.mp.  
34. intellect$ disable$.mp.  
35. mental retard$.mp.  
36. mental$ handicap$.mp.  
37. mental$ handicap$.mp.  
38. intellect$ impair$.mp.  
39. mental$ deficien$.mp.  
40. challenging behavi$.mp.  
41. self injur$.mp.  
42. inappropriate sex$  
43. behaviour.mp.  
44. violence.mp.  
45. anger.mp.  

| PsychInfo         | 1. DE = [(learning disorders) or (cognitive impairment) or (mental retardation)] or (mild mental retardation) or (severe mental retardation) (quality of life) or (client satisfaction) or (treatment outcomes)]  
2. DE = [(quality of life) or (client satisfaction) or (treatment outcomes)]  
3. DE = [(learning disorders) or (cognitive impairment) or (mental retardation)] or (mild mental retardation) or (severe mental retardation) (quality of life) or (client satisfaction) or (treatment outcomes)]  

| CINAHL            | MH mental retardation or MH developmental disabilities or MH outcome assessment or MH disability evaluation outcome and quality of life (outcome and quality of life) and (S1 and S2) (outcome and quality of life) and (S2 and S3)  

2. Narrowed search criteria: [(keywords: outcome)] or [(keywords: quality and keywords: of and keywords: life)]  


Accepted 6 April 2011