Confirmatory factor analysis of the quality of life scale and new proposed factor structure for the quality of life scale-revised

Kim T. Mueser [a,*], Min Kim [b], Jean Addington [c], Susan R. McGurk [a], Sarah I. Pratt [d], Donald E. Addington [c]

[a] Center for Psychiatric Rehabilitation, Departments of Occupational Therapy, Psychology, and Psychiatry, Boston University, Boston, MA, USA
[b] Center for Psychiatric Rehabilitation, Boston University, Boston, MA, USA
[c] Department of Psychiatry, University of Calgary, Calgary, Canada
[d] Department of Psychiatry, the Geisel School of Medicine at Dartmouth, Hanover, NH, USA; the Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth College, Hanover, NH, USA

ARTICLE INFO

ABSTRACT

The Quality of Life Scale (QLS) is a frequently used semistructured interview for the assessment of functional outcomes in schizophrenia. Despite the use of the QLS for over 30 years, the original 4-factor structure of the instrument (Interpersonal Relations, Instrumental Role, Intrapsychic Foundations, and Common Objects and Activities) has not been rigorously examined. Exploratory factor analyses (EFAs) and confirmatory factor analyses (CFAs) were used to evaluate the factor structure of the QLS in two independent datasets, including a mixed diagnostic sample of multi-episode participants ($N = 247$), and a sample of individuals with a first episode of psychosis ($N = 337$). A CFA with the first dataset indicated a poor fit for the 4-factor model of the QLS. Subsequent EFAs on this dataset led to a more promising 3-factor solution including 16/21 of the QLS items, which were similar to the first 3 of 4 factors originally proposed for the QLS. CFAs on the same dataset indicated that the 3-factor model for the QLS-Revised (QLS-R) fit the data well. This factor structure was evaluated with the second dataset using CFA and was also found to be fit the data well. The results support the robustness of the 3-factor model of the QLS-R in schizophrenia and mixed diagnostic samples. Future research should evaluate the validity of the 3-factor model of the QLS-R, and consider the merits of changing the name of the Intrapsychic Foundations factor to Motivation.

1. Introduction

The Quality of Life Scale (QLS) was developed over 30 years ago initially for the purpose of measuring the deficit syndrome (Heinrichs et al., 1984), a proposed subtype of schizophrenia characterized by persistent, primary negative symptoms (Carpenter et al., 1988; Kirkpatrick and Galderisi, 2008). The QLS is a semi-structured interview designed to assess four different areas of psychosocial adjustment, including Interpersonal Relations, Instrumental Role (e.g., work, school, homemaker), Intrapsychic Foundations (e.g., motivation, sense of purpose), and Common Objects and Activities (e.g., owning a watch, use of public transportation). Although the QLS was replaced by the Deficit Syndrome Scale as a measure of the deficit syndrome (Kimhy et al., 2006; Kirkpatrick et al., 1989), it has been widely used as a measure of psychosocial functioning in the intervening years since its publication (Bradley et al., 2006; Chou et al., 2012; Norman et al., 2000; Rabinowitz et al., 2012).

Numerous studies have supported the validity of the QLS as a measure of psychosocial functioning in schizophrenia and other severe mental illnesses (SMI) (Ascher-Svanum et al., 2013; Bellack et al., 1990; Faries et al., 2012; Thwin et al., 2013). Given the lack of a "gold standard" measure of psychosocial functioning in schizophrenia, the QLS has often been used in treatment studies. The QLS has enjoyed particularly widespread use in research on the treatment of prodromal psychosis states (Kim et al., 2013; McFarlane et al., 2012) and people recovering from a first episode of psychosis (Baksheev et al., 2012; Grootens et al., 2011; McEvoy et al., 2007; Perkins et al., 2004; Robinson et al., 2010). Despite the popularity of the QLS as an outcome measure in schizophrenia research, the factor structure of the QLS has not been empirically evaluated. Examination of the underlying factor structure of an instrument can improve the precision of the measurement of the central domains, as well as lead to refinements in the instrument that further increase its validity and sensitivity to change. This article describes the evaluation of the factor structure of the QLS based

 * Corresponding author at: Center for Psychiatric Rehabilitation, Boston University, 940 Commonwealth Ave. West, Boston, MA 02215, USA.
   E-mail address: mueser@bu.edu (K.T. Mueser).

on several study samples of people with schizophrenia-spectrum disorders and other severe mental illnesses.

2. Methods

The analyses were based on data from four de-identified treatment studies of people with schizophrenia and other severe mental illnesses. All of the studies received IRB approval from their corresponding institutions. For the purposes of the present study, two datasets were formed, in order to have a sufficient sample size to conduct confirmatory factor analysis, one based on three studies conducted at the Dartmouth Psychiatric Research Center and Center for Psychiatric Rehabilitation (Dataset #1, $N = 247$)) and the other based on a large study that was conducted at the University of Calgary (Dataset #2, $N = 337$). Demographic and clinical characteristics of the study samples are summarized in Table 1.

2.1. Dataset #1

Two of the three studies were two-site, randomized controlled trials evaluating a cognitive remediation program (Thinking Skills for Work) for persons enrolled in supported employment, with one site in Manchester, NH and the other site in Chicago, IL. The first study ($N = 107$) focused on people who had not benefitted from supported employment (McGurk et al., 2015), while the second study ($N = 93$), which is ongoing, is aimed at dismantling the critical components of the cognitive remediation program (McGurk, 2016). Inclusion/exclusion criteria for both studies were: 1) SMI according to States of NH or IL criteria; 2) unemployed and interested in competitive work; 3) enrolled in supported employment; 4) fluent in English; and 5) no evidence of traumatic brain injury or other medical condition with a major effect on brain functioning. An additional inclusion criterion for the first study was: participants had not benefitted from recent supported employment (i.e., had not worked competitively ≥3 months, or were fired or quit a job that lasted ≤3 months. Assessments used in these analyses included baseline, post-treatment (6–8 months later), and follow-ups 12–18 months and 24 months later.

The third study ($N = 47$) was an open pilot study of an individual adaptation of a group-based integrated psychosocial rehabilitation and healthcare intervention for older persons with severe mental illness (Mueser et al., 2010), conducted at two sites in NH. Inclusion/exclusion criteria for the study were: 1) SMI according to State of NH criteria; 2) age ≥ 50; 3) pervasive impairment across multiple areas of psychosocial functioning; 4) residing in the community; 5) no dementia or terminal illness with life expectancy ≤1 year; and 6) no cognitive impairment in moderate or worse range, as determined by Mini-Mental State Examination score >23 (Folstein et al., 1975). Assessments used in these analyses included baseline, post-treatment (approximately 12 months later), and a follow-up 3 months later.

For the purposes of analyses, assessments across the three studies were combined corresponding to four time periods: 1 (baseline), 2 (post-treatment), 3 (first follow-up), and 4 (second follow-up, Studies 1 and 2 only).

2.2. Dataset #2

This study ($N = 337$) took place in Calgary, Canada, and included participants recovering from a first episode of psychosis who had been engaged in a specialty treatment program (Addington and Addington, 2001). Inclusion/exclusion criteria for the study were: first episode of non-affective psychosis of any duration; three months or less of prior acceptable treatment; fluent in English; and no history of neurological disorders, head injury, or epilepsy (Larsen et al., 1996). The baseline assessment was used in these analyses.

2.3. Measures

Psychiatric diagnoses were based on the Structured Clinical Interview for DSM-IV (First et al., 1996), administered by trained clinical interviewers. Diagnoses were based on information obtained from the interviews, supplemented by chart review and information from informants (usually family members) when available.

The Quality of Life Scale (QLS) (Heinrichs et al., 1984) is a 21-item interview-based assessment that evaluates the adequacy of an individual's psychosocial functioning over the past month on 7-point Likert scales, ranging from 0 (poor) to 6 (good). The items on the QLS are divided into four a priori defined constructs or subscales, including: *Interpersonal Relations* (8 items tapping quantity and quality of social relationships), *Instrumental Role* (4 items tapping occupational, school, or parental functioning), *Intrapsychic Foundations* (7 items tapping core capacities believed to underlie effective interpersonal and instrumental functioning, such as motivation, curiosity, and sense of purpose), and *Common Objects and Activities* (2 items tapping possession of objects, such as owning a watch, and engagement in regular activities, such as shopping for food, thought to reflect active participation in the community). Prior to evaluating study participants, all interviewers were trained on the QLS. Training included didactic presentations, directed reading, receiving feedback on QLS interviews, and rating and discussion of videotaped QLS interviews, with interviewers achieving ICC ≥0.80 agreement on QLS subscales.

For all four studies, the QLS was administered as part of a more extended evaluation of clinical and psychosocial functioning, including a measure of psychopathology, either the Brief Psychiatric Rating Scale (BPRS) (Lukoff et al., 1986; Overall and Gorham, 1962) or the Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1987). Raters were instructed to include information obtained from the entire interview when rating each of the instruments.

2.4. Statistical analysis

Our general statistical approach was based on one we previously developed to evaluate the factor structure of the BPRS in two samples of persons with schizophrenia-spectrum disorders (Mueser et al., 1997). In that study, we used one dataset to examine the fit of previously proposed factor structures to the data using confirmatory factor analysis (CFA), to then conduct exploratory factor analyses (EFA) when a poor fit was found, and then to use CFA to evaluate and refine the fit of the new factor structure. In the second step we conducted several EFAs on different subsamples of the first dataset, varying with respect to gender and diagnosis, in order to identify a factor structure that would be maximally reliable across different dimensions of client heterogeneity. This new factor structure was then evaluated with CFA on the independent second dataset.

Similarly, in this study we first used CFA to evaluate whether the Time 1 QLS ratings from Dataset #1 fit the four-factor model described by Heinrichs et al. (1984). When this model was found to fit the data poorly, we conducted an EFA on the same data to identify an alternative factor structure, using maximum likelihood estimation and direct oblimin rotation for the final solution. Scree plots and communalities were examined to identify a more parsimonious solution, which was then replicated with EFAs across Times 2–4 on Dataset #1. We then performed CFAs on all four assessment points for Dataset #1 to evaluate the robustness of the factor solution over time. This model was then evaluated with CFA using the independent Dataset #2. All CFAs were conducted with AMOS version 21.0.0.

Model goodness of fit was evaluated using several indices, including the model χ^2 test, the normed fit index (NFI), the Tucker-Lewis index (TLI), the comparative fit index (CFI), and the root mean square error of approximation (RMSEA). The model χ^2 test compares the proposed factor structure to the null model with significant p values indicating inadequate model fit. However, the model χ^2 test is strongly influenced