MONASH University
Business and Economics

Centre for Health Economics

Research Paper 2011 (61)

Increasing the sensitivity of the AQoL inventory for the evaluation of interventions affecting mental health

Jeff Richardson
Foundation Director, Centre for Health Economics
Monash University

Gerald Elsworth
Assoc Prof (Hon) Public Health Innovation
Deakin Population Health Strategic Research Centre
Deakin University

Angelo Iezzi
Research Fellow, Centre for Health Economics
Monash University

Munir A Khan
Research Fellow, Centre for Health Economics
Monash University

Cathy Mihalopoulos
Senior Research Fellow, Health Economics Unit
Population Health Strategic Research Centre
Deakin University

Isaac Schweitzer
Healthscope Chair of Psychiatry
The University of Melbourne

Helen Herrman
Professor of Psychiatry, Centre for Youth Mental Health
The University of Melbourne

April 2011

Centre for Health Economics
ISSN 1833-1173
ISBN 1 921187 60 3

Correspondence:

Jeff Richardson
Centre for Health Economics
Faculty of Business and Economics
Monash University Vic 3800
Australia

Ph: +61 3 9905 0754 Fax: +61 3 9905 8344
Email: Jeffrey.Richardson@monash.edu

SYNOPSIS

This paper presents the rationale and psychometric analysis for extending the inventory of the Assessment of Quality of Life (AQoL)-6D instrument. The resulting AQoL-8D has an 8 dimensional, 35 item inventory with greater sensitivity in the domain of mental health.

The paper briefly reviews the existing QoL instruments used for economic evaluation of health programs. It outlines the steps adopted in developing the AQoL descriptive inventories and, specifically, the methods adopted for data collection and analysis for the AQoL-8D inventory.

Three instruments are presented. The first, PsyQoL, is a 22 item instrument which represents the best statistical fit for the measurement of mental health related quality of life. The second, PsyQoL-Brief is a reduced form instrument which is combined with AQoL-6D as the basis for the third instrument, the AQoL-8D. Psychometric properties of the first instrument are excellent and the second are good. The full AQoL-8D has satisfactory properties. Results from a comparison with the original AQoL-6D are reported. The mental health content of AQoL-8D is unique amongst MAU instruments and, along with other AQoL instruments, unique in its derivation from psychometric analysis. Its application to mental health patients and the public demonstrates its ability to discriminate between the groups with greater sensitivity than the previous AQoL-6D instrument.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Boxes

Increasing the sensitivity of the AQoL-6D for the evaluation of interventions affecting mental health

1 Introduction

Economic evaluation of health related interventions requires a comparison of the intervention costs and benefits. When these include the quality of life (QoL), benefits are increasingly measured using quality adjusted life years (QALYs). In this context, QoL is generally measured with a multi attribute utility instrument (MAUI), a multi-dimensional inventory (or descriptive system) with an associated formula for attaching preference weights to the items of the inventory. The purpose of the present article is to describe how the inventory of one such instrument, the Assessment of Quality of Life 6D (AQoL-6D) was enhanced to make it more suitable for use when a person's mental health is an important part of the evaluation.

Only a small number of MAUI exist and in recent years the field has been increasingly dominated by two instruments, the EQ-5D and Health Utility Index (HUI) (Brazier, Ratcliffe et al. 2007; Richardson, McKie et al. 2011). In Europe the SF-6D – a six dimensional reduced form of the SF-36 – is gaining popularity. The Quality of Wellbeing (QWB), the oldest instrument, is still commonly used in the USA. Other instruments exist such as the 15D and AQoL, but their use has been largely confined to the countries of their origin, Finland and Australia respectively. Only a few studies have compared results from 3 or more of these instruments. Authors of these have, without exception, found relatively limited association between instrument values and most have concluded that either the content of the instruments varies and/or that comparisons between scales should be undertaken with caution (Hawthorne, Richardson et al. 2001; Fryback, Palta et al. 2010; Khan and Richardson 2011).

The reason for focusing upon the mental health dimension of the AQoL-6D is twofold. First, poor mental health imposes a significant and rising burden of disease. Secondly, the explicit mental health content in the instruments cited above is very limited, (see Box 1). In principle the general items of an MAUI could be sensitive to mental health if the number of respondents ticking a lower category increased in exact proportion to the severity of each of the facets of a mental health problem. However this felicitous result has never been demonstrated and the more likely conclusion is that instruments are unequally sensitive to mental health states, implying unequal evaluation of problems involving these states. This tentative conclusion is supported by results from Böckerman et al. (2011) which indicate that even after standardising for two MAUI (EQ-5D, 15D) common psychiatric disorders decrease subjective wellbeing scores by 4 percentage points.

Box 1 Explicit mental health components of the major MAUI

EQ-5D	
'Anxiety Depression'	(1) Not anxious or depressed; *to* (3) Extremely anxious or depressed (Dolan, Gudex et al. 1995; Brooks, Rabin et al. 2005)
HUI 2	
'Emotion'	(1) Generally happy and free from worry; *to* (5) Extremely fretful angry irritable anxious or depressed, usually requiring hospitalisation or psychiatric institutional care (Torrance, Feeny et al. 1996)
HUI 3	
'Emotion'	(1) Happy and interested in life; *to* (5) Unhappy that life is not worthwhile (Feeny, Furlong et al. 2002)
SF-6D	(version 1: reduced from SF36)
'Mental Health'	(1) You feel tense or downhearted and low *none of the time; to* (5) You feel tense or downhearted *most of the time* (Brazier, Roberts et al. 2002)
SF-6D	(reduced from SF-12)
'Mental Health'	(1) You feel downhearted and low *none of the time; to* (5) You feel tense or downhearted *most of the time* (Brazier and Roberts 2004)
QWB	
Symptom list includes:	Spells of feeling upset, being depressed or of crying; *and*- Excessive worry or anxiety (Kaplan and Anderson 1996)
15D	
'Depression'	(1) I do not feel at all sad, melancholy or depressed; *to* (5) I feel extremely sad, melancholy or depressed
'Distress'	(1) I do not feel at all anxious stressed or nervous; *to* (5) I feel extremely anxious stressed or nervous (Sintonen and Pekurinen 1989)
AQoL-4D	
'Feeling'	(1) I do not feel anxious, worried or depressed; *to* (5) I am extremely anxious, worried or depressed (Hawthorne, Richardson et al. 1999)
AQoL-6D	
'Despair'	How often did you feel despair over the last 7 days (1) Never; *to* (5) All the time
'Worry'	How often did you feel worried (1) Never; *to* (5) All the time
'Sadness'	How often did you feel sad (1) Never; *to* (5) Nearly all the time
'Tranquillity'	Were you calm and tranquil or agitated (1) Always calm and tranquil; *to* (5) Always agitated (Richardson, Day et al. 2004)

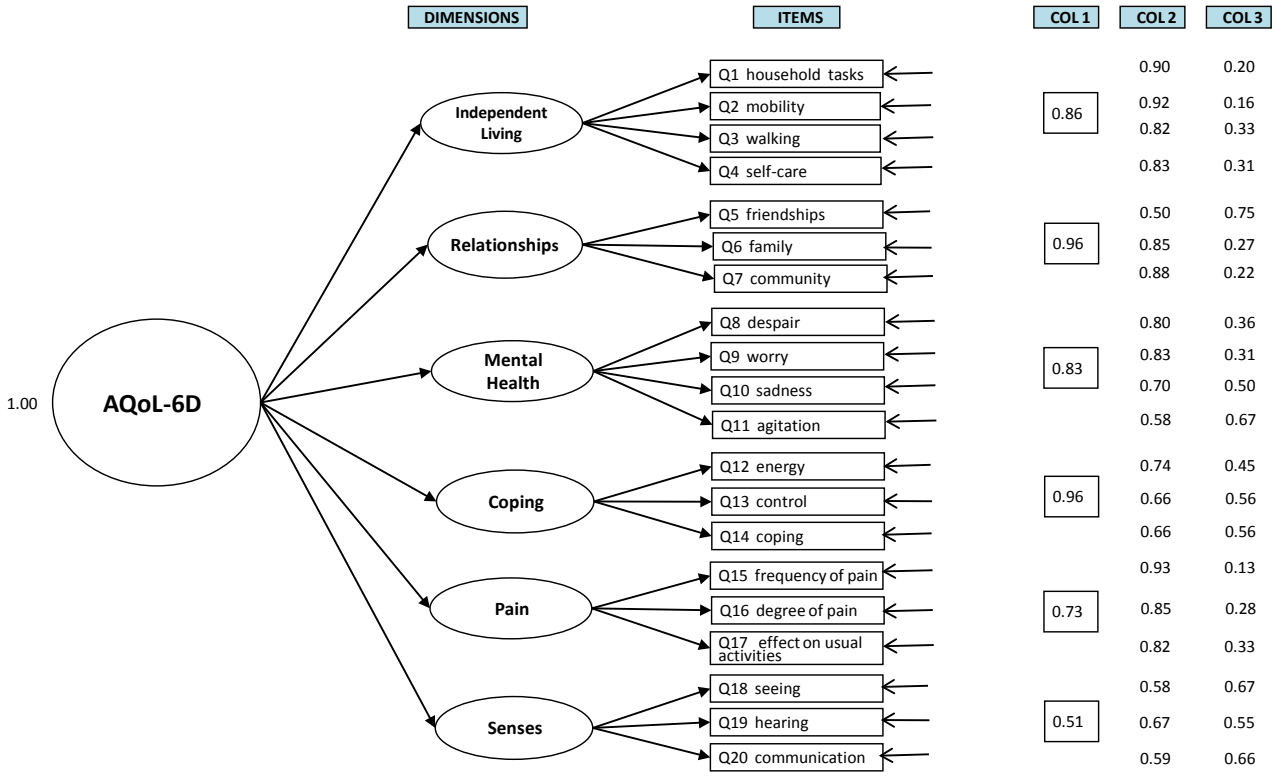
The broad approach to the construction of health related quality of life (HRQoL) instruments with content validity is discussed by a number of authors (Fayers and Machin 2000; Streiner and Norman 2003; O'Connor 2004). First, an overall concept (or theory) of health must be selected. This is operationalised by determining the dimensions of the concept and describing the postulated elements of each dimension with multiple items. These are initially analysed using expert input to ensure an adequate sampling of key elements and then edited for grammatical clarity and consistency. The surviving elements in the 'item bank' are administered to a representative sample of the relevant population to obtain the relationship between item responses. These are examined statistically to achieve item reduction while retaining the integrity

of the dimension content and the overall construct. In practice, this implies the interplay of statistical and substantive arguments. Items are excluded if the resulting scale is theoretically inconsistent irrespective of the statistical properties. Conversely, when there are strong *a priori* grounds for including items, they must also meet minimum statistical standards or be rejected.

The AQoL instruments (Box 2) were the first MAUI to adopt these broad procedures (Richardson, McKie et al. 2011). AQoL-4D (initially named 'AQoL' and published with 5 dimensions) commenced with a handicap based concept of QoL as it was hypothesised that this is more closely related to preferences than alternative concepts. There were few other constraints except the goal of instrument brevity. Four dimensions resulted: independent living, relationships, senses and mental health. AQoL-6D extended the model by adding coping and pain and additional items to independent living and mental health. Response categories were increased to extend 'normal' to 'excellent' health. The resulting structure and psychometric properties of the instrument are shown in Figure 1. The present analysis built upon, rather than replaced this work.

The present paper proceeds as follows. Section 2 below summarises the methods used and the rationale for the statistical procedures. Item and model selection are discussed in the following section and the characteristics of the model in Section 4. The concluding section discusses the problem of instrument comparison and specifically in the context of economic evaluation.

Figure 1 Structure of AQoL-6D



Box 2 AQoL instruments and related terminology

AQoL-8D
: The final 8 dimensional, 35 item instrument developed in this paper

AQoL-6D
: A 6 dimensional, 20 item instrument (Richardson, Day et al. 2004)

AQoL-4D
: The original 4 dimensional, 12 item instrument, initially called AQoL (Hawthorne, Richardson et al. 1999)

AQoL 8
: AQoL-4D less one item per dimension, estimated by interpolation 8 (Hawthorne, Richardson et al. 1999; Hawthorne 2009)

MAUI
: Multi attribute utility instrument

PsyQoL
: A 22 item inventory derived in this paper for mental health quality of life

PsyQoL-Brief
: A reduced form of PsyQoL which in combination with AQoL-6D is the basis of AQoL-8D

Content validity
: 'A judgement whether an instrument samples all the relevant or important content or domains' (Streiner and Norman 2003 p7).

Construct validity
: 'Linkage of a construct to some other attribute by a hypothesis or construct' (Streiner and Norman 2003 p8).

2. Methods

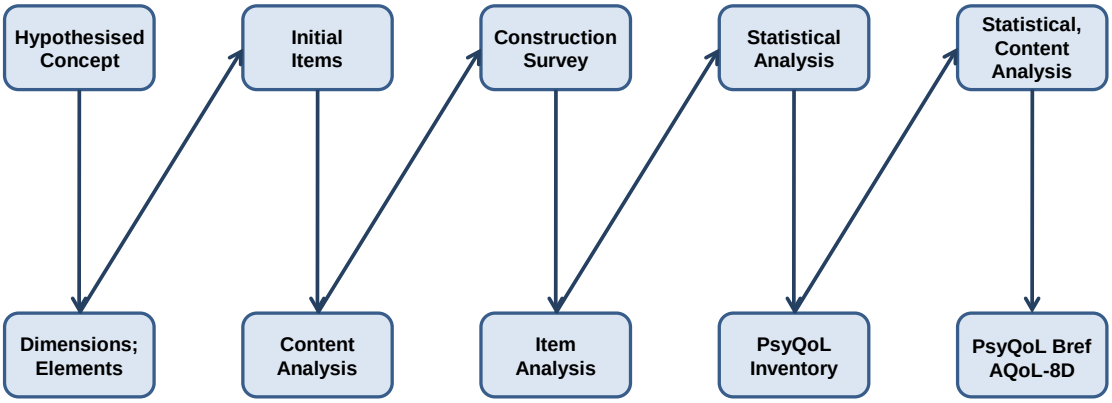
Figure 2 summarises the methods. As noted, AQoL-8D adopted the same concept of health – handicap – as the previous AQoL instruments. This basic conceptualisation was supplemented, when necessary, with elements of disability and impairment. The concept was operationalised by postulating dimensions of QoL elements and identifying or creating items which described these.

Initial Items: A literature search related to mental health quality of life was undertaken which identified several commonly used instruments. These are summarised in Table 1. Key items and elements from each instrument were collated and this provided the starting point for the development of the item bank. Four focus groups were conducted with 29 participants consisting of mental health patients and carers. Transcripts were examined to identify new issues and to guide the grammatical construction of new items.

Table 1 Instrument domains

	1	2	3	4	5	6	7	8	9	10
	Lehman	Wisconsin	Lancashire	Bigelow	Endicott	McSad	SchizoQLS	UCDI	Becker	AQoL-6D
SOCIAL	Social relations	Social relations	Social relations		Social relations			Community role	Social relations	Social relations
	Family	Family relations	Family relations					Family/ friends		
LIVING	Living Situation	Daily living	Living situation		Household duties			Living environment	Daily living	Daily living
	Work/ School	Activities and work	Work/ education	Role function	School/ Work	Role function		Work/ school	Occupational activities	
	Daily Activities				General activities			Daily activities		
HEALTH	Health in general	Symptoms/ outlook	Physical Health		Physical health	Sleep/ physiology	Symptoms and side effects		Physical health	
		Medicine/ illness							Symptoms	
										Pain
MENTAL HEALTH	Global Rating		General Wellbeing							
		Psychological Wellbeing	Mental Health				Motivation and energy		Psychological well-being	Mental Health
		Satisfaction							Satisfaction level	
			Self-concept		Affect	Affect/ Behaviour	Psychosocial			
		Goal attainment	Religion						Goal attainment	Coping
						Cognition				
OTHER	Finances	Finances	Finances					Income	Finances	
	Legal and Safety		Legal and Safety					Legal		
		Alcohol/ drugs								
			Leisure		Leisure					

Figure 2 Construction of the descriptive system



Content Analysis: The initial items were subject to linguistic and content analyses by the research team which consisted of two psychiatrics, a psychologist, a counsellor, a psychometrician and a health economist. Items were selected from the initial pool to obtain multi item coverage of each key element. The resulting items were provided to independent academic clinicians for feedback which led to further item modification.

Construction Survey: The construction survey administered the selected items to a sample of mental health patients and to a stratified sample of the Melbourne public selected to represent the demographic characteristics of the Australia population. The public sample was drawn from a computer readable phone directory, using a stratified, clustered two-stage design, similar to Hawthorne et al.'s (1999) procedures in the AQoL-4D validation study as well as the electoral roll. Based on the Australian Bureau of Statistics Socio-Economic Index for Areas (SEIFA) scores, postcodes were the primary sampling unit, with selection probability proportionate to population size (to reduce the effect of socio-economic confounding). From these postcode areas, telephone subscribers (18 years+) were sampled. Subscribers were contacted by letter and subsequently by telephone. The use of post-codes as the primary sampling unit meant that informants would be tightly clustered, minimizing the travel costs.

A wide range of psychiatric patients with mood and psychotic disorders were accessed via a number of channels. Mental Health Services (including St Vincent's Hospital Mental Health Unit) were used to recruit subjects. Case-managers and treating clinicians were approached to ensure people were well enough to participate in the study. Other private organisations such as The Melbourne Clinic and treatment providers for posttraumatic stress disorder were also approached to assist in the recruitment of people with non-psychotic disorders. Informed consent was obtained from all potential subjects.

Past experience indicated that to obtain a satisfactory response rate it would be necessary to make a payment in compensation for any out of pocket costs, such as travel costs. Interviews for the public took place at an AQoL research office. Patients were all interviewed in their hospital or treatment service site.

2.1 Statistical Approach

Item selection and scale validation were conducted in the tradition of classical test theory using unrestricted and restricted factor analyses. These correspond with the less descriptively accurate terms 'exploratory' and 'confirmatory' factor analyses. There is a distinction between the approximation of independent clusters in unrestricted factor analysis (which most analytic rotation methods tend to do) and Thurstone's original definition of a simple structure, which allowed each variable to be factorially complex, with large loadings on two or more factors. Two rotation approaches are designed to provide an approximation to Thurstonian simple structure. They are, first, the use of Cureton-Mulaik weights for row standardisation (as an alternative to the much more widely used Kaiser row standardisation) in combination with, for example, Varimax or Quartermax rotation, and secondly, the use of Yate's Geomin rotation criterion.

The combined use of unrestricted and restricted factor analysis for the validation and refinement of summed scales is recommended by, McDonald (2005) in the following passage:

> There is a case for accompanying a restricted, fully confirmatory factor analysis, in which the pattern of the entire factor loading matrix has been specified on substantive grounds, by a parallel unrestricted factor analysis that is also confirmatory in intention. This allows us to see if anything has been missed by the former. If the two analyses are essentially consistent, the overall weight of evidence is increased. If there are points of inconsistency, these may lead – at least as safely as the use of modification indices – to modified hypotheses for further test. Hopefully it will not lead to the unthinking freeing up of zero loadings in the restricted factor analysis that appear to correspond to salient loadings in the unrestricted analysis, but will instead stimulate further careful conceptualisation (McDonald, 2005, p. 171).

Unrestricted factor analysis may, additionally, be used to assess the number of factors needed to provide an adequate explanation of the associations in the data set, irrespective of the particular relations between indicators and factors. This assessment is the recommended first step in Mulaik's four-step approach to developing and validating a full structural equation model (Mulaik and Millsap 2000).

> This model (the unrestricted model) tests the hypothesis that a common-factor model fits the covariance matrix among the observed variables for a specified number of common factors *k* without confounding that issue with "measurement" issues about the specific relations of indicators to latent variables or structural relations between latent variables. The unrestricted model is effectively an exploratory common factor analysis model for the same number of factors as the number of latent variables of the structural equation model. It should be estimated with the same method of estimation as used to test the structural equation model …. It will be important to obtain the chi-square goodness-of-fit test from this analysis to be used in constructing goodness-of-fit indexes. It is essential to understand that the test of this model is a test of constraints implicitly in the fully specified structural equation model. Unacceptable fit obtained for the unrestricted model would also indicate unacceptable fit for the structural equation model (Mulaik and Millsap, 2000, p 43).

McDonald (2005) recommends the program Comprehensive Exploratory Factor Analysis (CEFA) for unrestricted factor analyses (Browne, Cudeck et al. 1998). Importantly the program also provides an ordinary least squares (OLS) estimator for use with polychoric correlations for the analysis of ordinal item response data.

For these reasons the unrestricted factor analyses of AQoL-8D item responses were carried out with CEFA. Using polychoric correlations, an initial OLS extraction was followed by two

contrasting oblique rotations: (a) Geomin, and (b) CF-Quartimax rotation (with traditional Kaiser row weights). [1]Rotation to the Geomin criterion has been shown to approximate Thurstonian simple structure by providing (rarely obtained) good orthogonal and oblique solutions to Thurstone's 26 variable Box problem (Browne, Cudeck et al. 2008 p 42). This criterion should thus enable the investigation of factorial complexity in the AQoL-8D items. In contrast, the CF-Quartimax solution with Kaiser weights should increase the likelihood of obtaining relatively independent item clusters.

Weighted Least Squares estimation based on polychoric correlations among the items is the theoretical method of choice for the restricted factor analysis of ordinal data (Jöreskog 2001). This method, however, requires large samples, particularly if the number of variables in the analysis is also large, and there is an emerging consensus that the requisite sample size may be very large indeed. Possible alternative approaches are: (a) various Diagonally Weighted Least Squares approaches (where only the diagonal elements of the appropriate weight matrix are used); and (b) Maximum Likelihood estimation with an appropriate weight matrix to correct for non-normal data distributions and robust (Satorra-Bentler) chi-square and associated 'close fit' statistics. The WLSMV approach in Mplus (version 2.13) was used for the present analyses. Here only the diagonal elements of the weight matrix are used in the estimation while the full weight matrix is used to compute standard errors and X^2 (Hancock and Mueller 2006 p 293).

Unstandardised and standardised factor loadings, an estimate of the variance in the measured variables explained by the latent variable (R^2) and associated standard errors are provided in Mplus 2 together with fit statistics (X^2, CFI – Comparative Fit Index, TLI – Tucker-Lewis Index, RMSEA – Root Mean Square Error of Approximation, and WRMR – Weighted Root Mean Square Residual). Recommended cut-off values for the tests of 'close fit' used in this report (CFI, TLI, RMSEA and WRMR) are based on recommendations from Yu (2002)) CFI ≥0.96; TLI ≥0.95; RMSEA ≤0.05; WRMR ≤1.0). Additionally, it is typically assumed that a value of the RMSEA that is ≤0.08 indicates a 'reasonable' fit (Browne and Cudeck 1993 pp 136-162).[2] Mplus also provides statistics that can be used to facilitate model improvement by suggesting fixed parameters (eg in the case of single-factor models, correlations among residuals) that might be freely estimated. In Mplus 2 with WLSMV these statistics include standardised residuals and derivatives (unscaled modification indices, useful when indicators are measured on similar scales[3].

3 Results

Three instruments were constructed. The first, PsyQoL, was specific to mental health. The second, PsyQoL Brief was a reduced form of PsyQoL which could be added to the AQoL-6D. The third, AQoL-8D was the best fit combination of AQoL-Brief and AQoL-6D which retained AQoL-6D

[1] CF-Quartimax is equivalent to direct quartimin in oblique as used here (Browne 2001).

[2] The Chi-square test of 'exact-fit' of the specified model to the data is directly sensitive to sample size and is typically regarded as much too stringent for practical purposes with large samples and complex models. Consequently it was not used as a specific criterion for model selection/rejection. Chi-square values are only quoted in those (rare) cases where the test was non-significant.

[3] Results were later verified and extended with Mplus Version 6 which also provides scaled modification indices for WLSMV. These can be interpreted as the increase in X2 that would result if the particular parameter were to be freely estimated and were used in conjunction with the standardised residuals to check previous modifications based on derivatives.

intact. The final model was a compromise between statistical criteria, the retention of items for theoretical reasons and the need for instrument brevity.

Item Bank: The initial item bank consisted of 250 items. As a result of linguistic and content analyses, it was reduced to 92 new items, the items from AQoL-4D and 6D, the K10 (a psychological distress scale) and one item relating to self assessed health, a total of 135 items.

Construction Survey: A total of 1503 letters were posted to the public. Subsequent analysis revealed 684 to be 'out of scope' (wrong address, moved, wrong telephone number, wrong age, or deceased). An additional 323 did not, as requested, make phone contact; 294 refused to participate and 202 provided useable results, a response rate of 40.6% of eligible contacts. Five sets of answers could not be used. Suitable patients were selected by doctors and carers. While all patients selected participated, the manner of approach precludes calculation of a meaningful response rate.

The construction survey was successfully administered by mail and by personal interview to 716 individuals, 514 patients (interviews) and 202 members of the public (mail). The larger number of patients was included in this survey as its primary purpose was to observe the relationship between item responses of individuals and, for the new module, it was necessary for individuals to have experienced the health states of interest. All data were screened for inconsistency and missing data.

Table 2 indicates that the overall age distribution of the sample population was satisfactory but, reflecting the difficulty in obtaining responses from the public, underrepresented young males. However this was offset by a disproportionate number of young males amongst the patient population. Similarly, from Table 3 the distribution of results across SEIFA groups was relatively uniform except for the larger number in SEIFA group 5. The missing SEIFA were due to mental health patients having no fixed address.

Table 2 Respondents to the construction survey by age and gender

Age	Public			Patient			
	Male	Female	Total	Male	Female	Total	Grand Total
18-24	1	6	7	109	42	151	158
25-34	5	21	26	60	64	124	150
35-44	13	25	38	51	63	114	152
45-54	24	32	56	41	52	93	149
55-65	28	42	70	11	21	32	102
	71	126	197	272	242	514	711

Table 3 Respondents to the construction survey by SEIFA group of residence*

SEIFA	Male	Female	Total	Male	Female	Total	Grand Total
1	14	25	39	47	25	2	111
2	10	24	34	48	28	76	110
3	24	27	51	24	22	46	97
4	13	21	34	34	32	66	100
5	10	29	39	75	89	164	203
	71	126	197	228	196	424	621

SEIFA Missing = 90

3.1 Construct validation and item selection

Unrestricted Factor Analysis: The eigenvalues of the matrix of polychoric correlations prior to factor extraction and rotation are shown in the form of a scree plot in Figure 3. This indicates that there are 9 factors with an eigenvalue greater than 1.0. There is also a sharp discontinuity between factors 7 and 8. The scree slope, however, is ambiguous, partly due to the discontinuity. There appear to be at least two identifiable 'elbows' in the plot, one suggesting 14 factors, the other 7. Thus these traditional rules of thumb suggest that between 7 and 9 factors might provide an optimal factor structure, but that the possibility of a greater number of factors (perhaps up to 14) should be investigated. The RMSEA for the 7 factor solution was 0.080 (90% confidence intervals 0.079 - 0.081) while the RMSEA for the 8 factor solution was 0.077 (90% confidence intervals 0.076 - 0.078). A value for the RMSEA ≤ 0.08 is typically regarded as indicating a 'reasonable' model fit. Both Geomin and CF-Quartimax rotated solutions were obtained for between 7 and 14 factors. The 8 factor CF-Quartimax solution was chosen as providing the most interpretable delineation of item clusters. These were identified as: *Financial Management*; *Happiness*; *Anger and Hatred*; *Pain* and (lack of) *Ease with Normal Physical Activities*; *Social Relationships* (including sexual relationships); *Self Worth*; *Loneliness and Social Isolation*; and *Anxiety* (including restlessness and worry). The factor loadings for the 8-factor CF-Quartimax solution are shown in Table 4. The table is arranged to highlight the relationship between the factor solution and the structure anticipated by the research team. Primary factor loadings are in italics.

Figure 3 Scree Plot of Eigen Values of Polychoric Correlations Among 92 'PsyQoL' Items

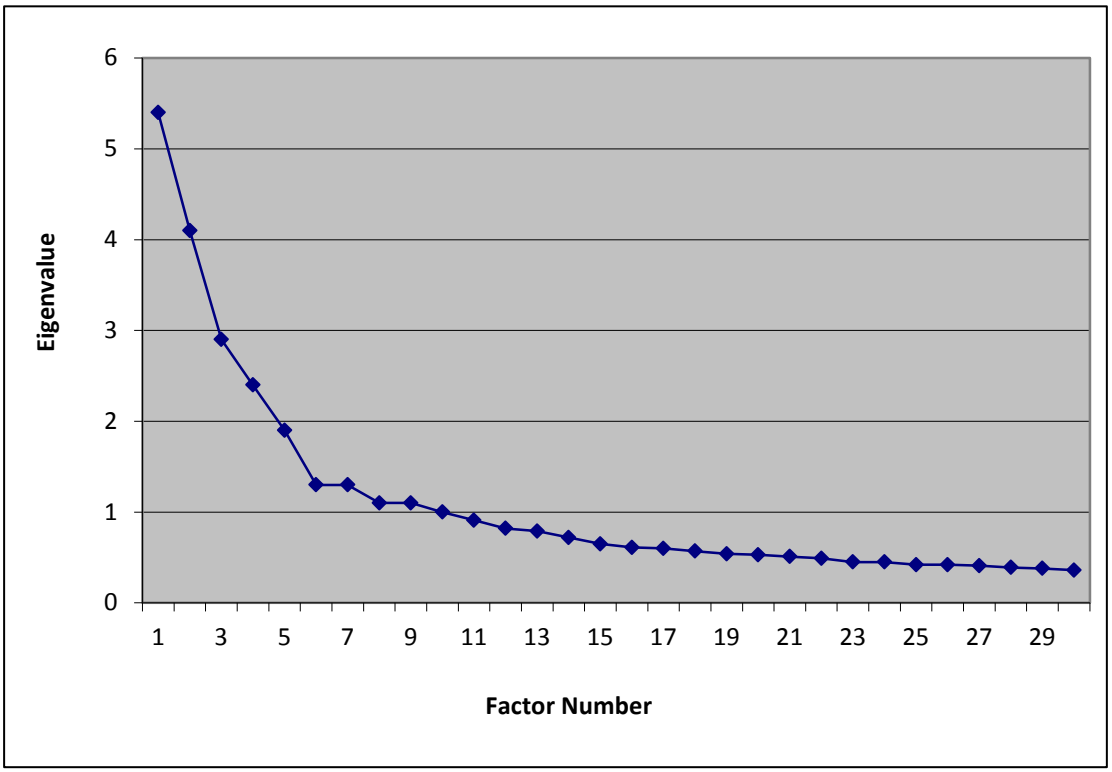


Table 4 Unrestricted Factor Analysis of 92 'PsyQoL' Items - Eight-Factor Oblique 'CF-Quartimax' Solution with Kaiser Weights

Eight-Factor Oblique 'CF-Quartimax' Solution – Kaiser Weights

		F 1	F 2	F 3	F 4	F 5	F 6	F 7	F 8
		Financ'l M'ment	Happiness	Anger & Hatred	Pain	Relat's (Family & Friends)	Self Esteem/ Worth	Lonel's & Social Isolat'n	Restles/ Worry/ Anxiety
	Independent Living								
Var1	How difficult is it for you to join in activities?	0.09	0.14	0.02	0.27	0.10	0.18	0.31	0.01
Var2	Are you able to take care of yourself?	-0.23	-0.03	-0.03	-0.13	-0.05	0.00	-0.35	-0.03
Var3	Are you able to do physical activities that are normal for your age?	0.25	0.14	-0.11	0.44	0.04	-0.06	0.25	-0.02
Var4	How easily can you manage jobs around the house (eg cooking, cleaning or gardening)?	0.23	0.21	0.02	0.34	0.02	-0.04	0.30	-0.02
	Social & Family								
Var5	Your close relationships (family and friends) are:	0.04	-0.03	0.08	0.03	0.76	0.06	-0.02	-0.03
Var6	How much do you enjoy your close relationships (family and friends)?	0.05	0.03	0.07	-0.01	0.76	0.10	-0.10	-0.07
Var7	Your close relationships (family and friends) are:	0.00	-0.03	0.07	0.00	0.81	0.07	0.01	-0.05
Var8	How much of a burden do you feel you are to other people?	0.23	0.00	0.08	0.13	-0.02	0.22	0.32	0.19
Var9	How lonely do you feel?	-0.05	0.05	0.10	0.09	0.26	0.09	0.45	0.13
Var10	How often do you feel lonely?	0.00	0.07	0.09	0.07	0.28	0.07	0.38	0.19
Var11	How often do you feel socially isolated?	-0.02	-0.07	0.07	0.11	0.28	0.17	0.54	0.04
Var12	How often do you feel socially excluded or left out?	-0.01	-0.12	0.05	0.20	0.27	0.26	0.42	0.00
Var13	How much do you feel socially excluded?	0.09	-0.01	0.04	0.22	0.19	0.11	0.44	-0.04
Var14	How satisfied are you with your friendships?	0.04	-0.08	0.02	0.05	0.65	0.13	0.13	-0.03
Var15	How much do you enjoy being with your friends?	0.09	0.11	0.05	0.06	0.47	0.18	0.08	-0.16
Var16	Thinking about your sexual relationships, whether you have any or not, are you:	0.06	0.11	-0.01	0.00	0.65	-0.13	0.05	0.13

Var17	Are you satisfied with your close and intimate relationships (including any sexual relationships)?	-0.01	0.01	-0.01	0.03	0.76	-0.09	0.06	0.11
Var18	How satisfied are you with your close and intimate relationships (including any sexual relationships)?	0.03	0.03	0.02	0.03	0.84	-0.10	-0.01	0.07
	Mental Health								
Var19	Generally you feel: Depressed etc.	0.05	0.22	0.20	0.14	-0.05	0.13	0.29	0.27
Var20	How often do you feel depressed?	0.07	0.21	0.25	0.07	-0.06	0.14	0.26	0.30
Var21	Generally you feel: Depressed etc.	0.04	0.37	0.18	0.07	0.07	0.20	0.11	0.21
Var22	Feelings of depression interfere with your life:	0.15	0.19	0.15	0.14	-0.02	0.13	0.23	0.29
Var23	How often do you have trouble sleeping?	0.03	0.16	0.08	0.28	0.13	0.05	-0.07	0.31
Var24	Most mornings, when you wake up, how do you feel?	-0.01	0.44	0.09	0.14	-0.01	0.05	0.05	0.19
Var25	How much does difficulty with sleeping bother you?	0.01	0.17	0.14	0.32	0.08	-0.06	0.00	0.34
Var26	How often do you feel restless?	0.13	0.05	0.17	0.06	0.18	-0.10	0.18	0.44
Var27	How much does restlessness disturb you?	0.16	0.08	0.11	0.12	0.15	-0.08	0.13	0.38
Var28	How content are you with your life?	0.11	0.37	0.06	0.03	0.23	0.19	0.16	0.10
Var29	How often do you feel calm?	0.22	0.11	0.12	0.08	0.08	0.17	-0.03	0.35
Var30	How often do you feel hopeless?	0.15	0.08	0.18	0.09	0.01	0.31	0.25	0.21
Var31	How confident do you feel about managing your money?	0.88	-0.05	0.02	-0.05	0.05	0.08	-0.08	-0.05
Var32	Managing your finance leads to: anxiety etc?	0.85	0.00	0.06	0.05	-0.03	-0.03	-0.04	0.05
Var33	Managing your finance makes you anxious:	0.83	-0.02	-0.01	0.01	0.06	0.00	-0.05	0.07
Var34	How often does anxiety interfere with your daily life?	0.26	0.01	0.02	0.17	0.13	0.17	0.03	0.40
Var35	How often do you feel anxious or nervous?	0.15	0.02	0.12	0.15	0.07	0.23	0.07	0.44
Var36	How anxious do you generally feel?	0.14	0.09	0.07	0.16	0.08	0.18	0.13	0.40
Var37	How often do you feel afraid?	0.14	-0.23	0.06	0.09	0.13	0.28	0.10	0.44
Var38	Thinking about fear, generally you feel?	0.11	-0.15	0.09	0.14	0.12	0.20	0.15	0.37
Var39	How well do you feel you handle stress?	0.25	0.18	0.04	0.06	0.05	0.21	0.08	0.24
Var40	How often do you feel worried?	0.25	0.13	0.15	0.04	0.00	0.20	0.04	0.37

Var41	How worried do you generally feel?	0.22	0.10	0.05	0.15	0.07	0.10	0.12	0.36
Var42	How often do you feel mixed up or confused?	0.26	-0.09	0.20	0.08	0.06	0.13	0.13	0.32
Var43	Are you confused by day to day bill paying?	0.63	-0.09	0.11	0.17	-0.01	-0.03	0.19	-0.07
Var44	How often do you feel hatred for others?	0.08	-0.12	0.78	0.03	0.05	0.11	0.05	-0.14
Var45	How much hatred do you feel for others?	0.04	-0.09	0.76	0.01	0.10	0.07	0.11	-0.19
Var46	How often do you feel angry?	-0.07	0.04	0.84	0.04	0.00	0.02	-0.03	0.11
Var47	How much anger do you feel?	-0.03	-0.01	0.90	0.01	0.02	-0.01	-0.04	0.06
Var48	How much anger and hatred do you feel?	-0.04	-0.01	0.91	0.02	-0.04	0.01	0.09	-0.01
Var49	How often do you feel anger and hatred?	-0.01	0.00	0.92	-0.01	-0.04	0.00	-0.01	0.05
Var50	Do you want to hit out at people or objects?	0.08	0.05	0.82	-0.02	-0.03	0.01	-0.10	-0.02
Var51	How often do you get angry and become physically or verbally aggressive?	0.07	0.01	0.71	0.08	0.01	-0.14	-0.02	-0.03
Var52	How much irritability or anger do you generally feel?	-0.03	0.10	0.77	0.00	0.09	-0.08	-0.02	0.16
Var53	Do you ever think of / feel like hurting yourself?	0.04	0.11	0.29	0.05	-0.01	0.27	0.18	0.21
	Coping								
Var54	Life's day by day problems, eg paying bills, managing house, coping with people:	0.34	0.29	0.09	0.11	0.04	0.10	0.04	0.07
Var55	How much control of your life do you feel you have?	-0.23	-0.11	-0.08	0.06	-0.01	-0.06	-0.42	-0.06
Var56	When problems arise, you can find a strategy to deal with them:	0.29	0.15	0.01	0.07	-0.01	0.13	0.30	0.11
Var57	When problems arise, you deal with them: adequately/poorly?	0.36	0.10	0.06	0.05	0.06	0.26	0.12	0.14
Var58	When problems arise, you cope with them: easily etc.?	0.33	0.05	0.03	0.08	0.04	0.25	0.20	0.16
	Pain								
Var59	How much pain do you experience?	-0.04	-0.06	0.05	0.96	-0.02	0.04	-0.03	-0.06
Var60	How often do you feel pain?	-0.01	0.03	0.01	0.93	0.00	-0.04	-0.12	0.03
Var61	How much does pain interfere with your ordinary activities?	0.04	-0.01	0.01	0.93	0.00	-0.07	0.02	-0.06
Var62	How distressing is your pain usually?	-0.06	-0.06	0.05	0.88	0.01	0.02	0.00	0.09
	Happiness								
Var63	How enthusiastic do	0.06	0.43	0.04	0.18	0.06	0.28	0.13	-0.03

	you feel?								
Var64	How often do you feel happy?	0.10	0.38	0.12	0.07	0.23	0.19	0.07	0.06
Var65	How often do you have fun?	0.05	0.41	-0.02	0.11	0.32	0.17	0.12	-0.03
Var66	Do you feel joy?	0.04	0.41	0.04	0.07	0.30	0.26	0.04	-0.07
Var67	How much pleasure do you get from your usual activities?	0.06	0.40	0.04	0.05	0.24	0.13	0.19	0.03
Var68	How much pleasure do you get from your life?	0.08	0.42	0.10	-0.01	0.22	0.12	0.17	0.10
Var69	How often do you feel pleasure?	0.05	0.45	0.09	0.06	0.27	0.22	0.05	0.00
Var70	How often do you feel joy and pleasure in life?	0.08	0.43	0.05	0.06	0.28	0.20	0.10	-0.03
Var71	How satisfied are you with the way you manage bills and finances?	0.87	0.06	0.03	-0.04	0.03	-0.06	0.02	0.00
Var72	You feel fulfilled and satisfied with your life:	0.10	0.36	0.07	0.01	0.29	0.14	0.15	0.09
Var73	How satisfied are you with your life?	0.09	0.35	0.12	0.02	0.20	0.20	0.19	0.03
Var74	How fulfilling is your life?	0.03	0.37	0.08	0.03	0.23	0.21	0.19	0.06
Var75	Thinking about purpose and goals in your life, you are: satisfied etc.?	0.07	0.39	0.12	0.03	0.17	0.16	0.16	0.08
Var76	How often do you have a sense of purpose and fulfilment about your life?	0.09	0.37	0.10	0.08	0.14	0.28	0.11	0.01
Var77	Thinking about yourself, you feel you are: very important etc.	0.04	0.19	0.13	0.03	0.21	0.49	-0.01	-0.05
Var78	Your self esteem is:	0.11	0.22	0.04	0.08	0.02	0.52	0.07	0.13
Var79	You feel you have good self esteem:	0.12	0.17	0.07	0.08	0.07	0.48	0.09	0.12
	Self Worth								
Var80	How often do you feel good about yourself?	0.07	0.33	0.10	0.08	0.08	0.38	0.02	0.16
Var81	How often do you feel worthless?	0.11	0.08	0.17	0.07	0.02	0.39	0.22	0.17
Var82	How much confidence do you have in yourself?	0.18	0.16	0.02	-0.03	0.07	0.48	0.12	0.13
Var83	How often do you feel significant?	0.10	0.17	0.14	0.04	0.15	0.41	0.11	-0.02
Var84	How often do you feel inferior to others?	0.15	-0.07	0.07	0.01	0.10	0.53	0.06	0.20
Var85	How many good qualities do you think you have?	0.14	0.15	0.04	0.02	0.09	0.51	0.08	0.00
Var86	How often do you feel that you are a failure?	0.16	0.07	0.12	0.05	0.04	0.39	0.18	0.21
	Leisure								
Var87	How bored do you feel with life?	0.03	0.30	0.20	-0.08	0.09	0.00	0.48	0.09

Var88	You feel bored:	-0.02	0.17	0.12	-0.05	0.11	0.03	0.48	0.13
Var89	How often do you enjoy your leisure time?	0.17	0.41	0.00	0.11	0.26	0.08	0.04	0.01
Var90	How much do you enjoy your leisure time?	0.10	0.41	-0.04	0.13	0.28	0.09	0.10	-0.06
Var91	How do you feel about how you spend your time?	0.06	0.43	0.07	0.08	0.11	0.04	0.31	0.04
Var92	How satisfied are you with the amount of leisure time you get?	0.12	0.30	0.16	0.02	0.12	-0.03	0.00	0.10

The Geomin solutions tended to be dominated by a general factor that (in the 8 factor Geomin-rotated solution) contained the depression and anxiety items from the *General Mental Health* construct, items denoting social isolation and loneliness from the *Social and Family* dimension, as well as most of the *Happiness*, *Coping*, *Self Worth* and *Leisure* items. The other clearly delineated item clusters in the 8 factor Geomin solution identified the 'anger and hatred' items from *General Mental Health*, 'social and family ties and sexual relations' from the *Social and Family* construct, 'financial management' (from *General Mental Health*) and *Pain*. One possible interpretation of the Geomin structure is that the very broad 'general mental health' grouping of items might represent a higher-order factor that could be further sub-divided, while the other four clusters are those that contain strongly associated uni-factorial items that constitute unique and independent clusters.

The CF-Quartimax solution was a moderately satisfactory replication of the *a priori* constructs specified by the research team. Clear *Happiness* and *Self Worth* item clusters were identified. The 'social and family' group of items divided into a clearly delineated *Social (including Sexual) Relationships* item cluster and another (not so clearly delineated) cluster denoting *Loneliness and Social Isolation*. The *General Mental Health* dimension broke up into three more specific clusters: *Financial Management* (which was also clearly identified in the Geomin solution), *Anger and Hatred* (also clearly delineated in the Geomin solution) and *Anxiety*. There was not, however, a clearly delineated 'depression' item cluster; items denoting depression were associated with both the *Anxiety* and *Happiness* item clusters. Finally, three of the four items denoting 'independent living' were weakly associated with a clearly delineated cluster of items that tapped the extent and impact of pain.

Restricted Factor Analysis: Based on the prior decision to build upon AQoL-6D, and the unrestricted factor analysis the following seven independent constructs were hypothesised: *Anger and Hatred*; *Anxiety*; *Depression*; *Coping*; *Happiness*; *Relationships* and *Self Worth*. *Anger and Hatred*, *Anxiety* and *Depression* were, together, hypothesised to constitute a second-order factor of *General Mental Health*. Validation of the homogeneity of these constructs together with selection of a parsimonious item set for each that retained acceptable breadth was undertaken by iteratively testing an extensive series of one-factor congeneric[4] models for each item cluster.

[4] A congeneric model is one where a set of items or scales are specified to be loaded by a single latent variable. Both factor loadings and error variances are free to vary, but additional relationships between error variances (correlated residuals) are not allowed.

Anger and Hatred

A cluster of items denoting *Anger* was specified as a desirable sub-construct of *General Mental Health* by the principal research team and an *Anger and Hatred* cluster of items was clearly defined in both Geomin and CF-Quartimax unrestricted factor analyses. On the basis of the research team's *a priori* specification and the factor analysis results, items 45 to 53 in Table 4, were initially hypothesised to be associated with this construct. However the fit of a single-factor model to these items was not satisfactory (CFI = 0.98; TLI = 0.970; RMSEA = 0.15; WRMR = 1.89). A satisfactory fit was subsequently achieved by including three correlated residuals in the model. (CFI = 0.99; TLI = 0.99; RMSEA = 0.08; WRMR = 0.90). Items 44, 45 and 48 that contributed to the two largest correlated residuals were all items that included 'hatred' in the stem. Deleting these items (but retaining an item that referred to both 'hatred' and 'anger' in the stem) and Item 50 that contributed to a correlated residual with Item 51 resulted in an excellent fit to the data (CFI = 1.00; TLI = 1.0; RMSEA = 0.06; WRMR = 0.50). The resulting cluster comprised Items 46, 47, 49, 51, 52, and 53. These items were principally focussed on feelings of anger, but included connotations of hatred, aggression and self-harm. Standardised factor loadings from the six item model ranged from 0.64 to 0.93 and the internal consistency (Cronbach's alpha) of the resulting summated scale was 0.90.*Anxiety*

An item cluster denoting *Anxiety* was similarly specified by the research team and identified as an independent cluster in the CF-Quartimax unrestricted factor analysis solution. Informed by the results of the unrestricted factor analysis, items 26-27, 29, 34–36, 39 and 40–42 in Table 4, were initially specified to constitute this cluster. While all items were loaded >0.79 on the single factor the fit was not satisfactory (CFI = 0.99; TLI = 0.99; RMSEA = 0.13; WRMR = 1.53). Modification indices suggested that there was a clear sub-factor constituted by Items 26 and 27, both referring to 'restlessness' and they were deleted from the model specification. While the model fit improved substantially there remained another significant correlated residual between Items 40 and 41; both denoting worry. Deletion of each was trialled and a satisfactory model was obtained with the deletion of item 41 (CFI = 1.00; TLI = 1.00; RMSEA = 0.06; WRMR = 0.59). This final model consisted of Items 29, 34-36, 39, 40, and 42. Three of these items referred directly to anxiety while others connoted being calm (reverse scored), stress and worry. Subsequently, the research team included Item 26 'How often do you feel restless?' in the *General Mental Health* construct in the final version of the inventory. Returning this item to the model for *Anxiety* resulted in a satisfactory fit (CFI = 1.00; TLI = 1.00; RMSEA = 0.07; WRMR = 0.70). Loadings of all items were >0.80 (range 0.80-0.95) and Cronbach's alpha for the resulting scale was 0.94.

Depression

A distinct cluster of items connoting depression was not identified in the unrestricted factor analyses; all apparently relevant items were quite strongly associated with a more *General Mental Health* cluster in the Geomin rotation but were scattered in relation to the constructs identified in the CF-Quartimax rotation (principally associated with anxiety but also with social isolation and self worth). A sub-construct of *Depression* was, however, identified by the research team as a critical component of *General Mental Health* and the possibility of a distinctive and homogeneous item cluster was tested in a restricted analysis. Items 19 – 23 and 30 were included in the cluster. The one factor model with no modifications was a satisfactory fit to the data (CFI = 1.00; TLI = 1.00; RMSEA = 0.07; WRMR = 0.55). With one exception, all items were loaded >0.89 by the factor. The exception was Item 23 where the loading was 0.66. Cronbach's alpha for the six item scale was 0.94

General Mental Health

For substantive reasons the research team selected the following items from the *Anger and Hatred*, *Anxiety* and *Depression* scales to constitute a possible *General Mental Health* scale: 20 (*Depression*), 23 (*Depression*), 26 (*Anxiety*), 29 (*Anxiety*), 30 (*Depression*), 35 (*Anxiety*), 37, 46 (*Anger and Hatred*), 53 (*Anger and Hatred*). Item 37, connoting fear was added to incorporate an omitted but important facet of emotional response. The fit of the one-factor model to these items was reasonably good (CFI = 0.996, TLI = 0.995, RMSEA = 0.066; WRMR = 0.739) and the internal consistency of the scale (Cronbach's alpha) was 0.93. With three exceptions factor loadings exceeded 0.8. The exceptions were 23 sleeping, (0.66), 37 fear (0.76) and 46 anger (0.68).

Happiness

A cluster of items (63-79) largely connoting pleasure, joy, happiness, fun and enthusiasm (together with two items specifically denoting 'self-esteem') was initially hypothesised to form a factor provisionally named *Happiness*. The factor was also quite well delineated in the unrestricted factor analysis, although the self-esteem items and one referring to financial management were not associated with the factor, while four items connoting leisure originally hypothesised to be associated with the *Self-Worth* construct were, along with a small number of items that tapped other emotional responses.

An initial restricted factor analysis of variables 63-70, 72-77 and 89-92 did not fit well (CFI = 0.984; TLI = 0.982; RMSEA = 0.124; WRMR = 1.827). The largest modification index pointed to a correlation between the residuals of items 89 and 90 (both referring to the enjoyment of leisure time). Subsequent fitting of a series of models with correlated errors suggested that items 89-92 (all connoting 'leisure') constituted a separate factor. Deleting these items and a small number of others that yielded correlated errors resulted in a model comprising items 63-69, 72-77, 89 and 91 with improved but still not acceptable fit (CFI = 0.993; TLI = 0.991; RMSEA = 0.094; WRMR = 1.267).

Seven items were chosen for substantive reasons to represent the *Happiness* construct: 28, 63-65, 67, 69 and 89. Item 28 ('How content are you with your life') was originally classified as a *General Mental Health* item, but was clearly associated with *Happiness* in the unrestricted factor analysis. It was included to increase the breadth of the construct. Likewise item 89, 'How often do you enjoy your leisure time?' was also included. The resulting construct encompassed the feelings of contentment (with life), enthusiasm, happiness, having fun, pleasure and enjoyment of leisure time. The fit of the one-factor model to this item set was not completely satisfactory however (CFI = 0.998; TLI = 0.997, RMSEA = 0.084; WRMR = 0.703). The RMSEA exceeds the conventional 'bench-mark' for a 'reasonable' fit of 0.80. Inclusion of two correlated residuals, both involving item 28, improved the fit (RMSEA = 0.061). The resulting correlated residuals were both negative. Negative correlated residuals do not suggest presence of a subsidiary factor structure and it was decided to retain Item 28 in the scale to maintain breadth. Cronbach's alpha for the seven item scale was 0.94 and all factor loadings exceeded 0.82.

Self Worth

Items 80-86 were initially hypothesised to constitute a scale measuring a construct of *Self Worth*. All were located on the same factor in the unrestricted analysis together with three items that referred to self-esteem and feelings of importance (78-79 and 77 respectively). These were originally classified with the *Happiness* construct, together with one item from the *General Mental Health* construct which focussed on feelings of hopelessness. However this item, 30, was judged

to be outside the boundaries of a construct of self-worth and an initial one-factor model was fitted to items 77-86.

The initial model fit was not satisfactory (CFI = 0.990; TLI = 0.987; RMSEA = 0.127; WRMR = 1.476). Item 77 was found to be associated with the largest correlated residual (with Item 83) and was deleted. Fitting four correlated errors in the resulting model yielded a marginally satisfactory fit (CFI = 0.996; TLI = 0.995; RMSEA = 0.082; WRMR = 0.875).

Subsequent discussion in the research team suggested that a model consisting of items 81-82 and 84-86 when combined with item 8 ('How much of a burden do you feel you are to other people?') would provide a construct of adequate focus and breadth, encompassing the major components of self-worth. The fit of the one factor model was not acceptable (CFI = 0.996, TLI = 0.994, RMSEA = 0.098, WRMR = 0.835). The fit could be improved considerably be adding a correlated residual between Item 82 and Item 85 (RMSEA = 0.051, residual correlation = 0.32). This correlated error contrasted the two items that are worded in a 'positive' direction with the other four, which have a 'negative' connotation. There was not, however, any clear overlap in content between Items 82 and 85 apart from the general direction of item wording. Removing one or other of these items negated the 'correlated residual' problem, but the scale would then be reduced to five items and would be biased towards 'negative' items. On balance, it appeared prudent to retain the six items.

The internal consistency of the resulting 6 item scale (Cronbach's alpha) is 0.93.

Social and Family Relationships

The items (5-18) were initially postulated to form a construct for *Social and Family Relations.* However these fell into two discrete sub-sets in the unrestricted factor analysis: 'family and sexual relations' and 'social isolation and loneliness'. In the restricted factor analyses it was not possible to fit a satisfactory one-factor model that incorporated these two sub-sets (without modifications the fit statistics were CFI = 0.929, TLI = 0.916, RMSEA = 0.200, WRMR = 3.508). Indeed, there was strong evidence from the pattern of correlated residuals that a three-factor model that distinguished 'relationships with family and friends' from 'sexual relationships' and 'loneliness and social isolation' was required. After a number of iterations, it was concluded that the most satisfactory model that could be obtained was a three-factor solution incorporating six items representing the three sub-constructs specified above. The fit of the three-factor model was very good (CFI = 0.999, TLI = 0.999, RMSEA = 0.039, WRMR = 0.293). Additionally, the chi-square for this model was non-significant X^2= 12.332, 6 d.f. pr = 0.055. Correlations between the factors were: *Relationships (Family and Friends)* with *Sexual Relationships* = 0.80; *Relationships (Family and Friends)* with *Social Isolation* = 0.72; *Sexual Relations* with *Social Isolation* = 0.67.

The results for this item set posed a dilemma for scale construction. Including all 6 items in a single scale clearly violates the fundamental principle in psychometrics of 'local independence' (the principal that all systematic variance in a homogenous scale should be accounted for by the single latent variable). There were, however, insufficient items to represent each sub-set adequately as single scales and it was argued that the three item sub-sets represented constructs that were too specific for separate inclusion in a generic 'psychological quality of life' inventory. Further, attempting to circumvent the problem by retaining one item only for each sub-set would result in a scale that contained only three items with consequent low scale reliability.

The best available solution appeared to be to accept the composite six-item set as an 'aggregate' of three separate, but quite strongly related, constructs. The internal consistency of this aggregate scale (Cronbach's alpha) is 0.89.

Coping

The initial conception of psychological quality of life included a construct of 'coping', defined by items 54-58 in the questionnaire. A 'coping' factor was not, however, identified in the unrestricted factor analysis; the 'coping' items were split between an item cluster that was predominantly defined by items connoting satisfactory financial management and one encompassing feelings of loneliness, social isolation and boredom.

The research team deemed the ability to handle stress to be a necessary component of a satisfactory conception of coping despite its exclusion in the initial analysis. Item 39 was identified as the most suitable stress related item and was included in the scale. The fit of the one-factor model comprising Items 39 and 54-58 was, however, very poor (CFI = 0.932, TLI = 0.904, RMSEA = 0.280, WRMR = 5.181). Subsequent removal of item 54 (finding 'life's day to day problems … enjoyable', easy etc.) and item 55 ('How much control of your life do you feel you have?') resulted in a four-item set for which a one-factor model fitted very well (CFI = 1.000, TLI = 1.000, RMSEA = 0.016, WRMR = 0.184 – also Chi-square = 2.386, 2d.f. pr = 0.303). Cronbach's alpha for the resulting scale was 0.90.

3.2 The full measurement model for psychological quality of life (PsyQoL)

With the three facets of 'family and social relationships' separated the items and factors specified above, form a seven factor model. This fitted the data reasonably well (CFI = 0.989, TLI = 0.987, RMSEA = 0.058, WRMR = 1.132). The model is rigorously specified as each item was required to be loaded by only one factor (no 'cross-loadings' were allowed) and there were no correlated residuals. Factor loadings were in the following ranges: (i) *General Mental Health* 0.67-0.96; (ii) *Happiness* 0.82-0.94; (iii) *Self Worth* 0.81-0.92; (iv) *Relationships (Family and Friends)* 0.84-0.90*; (v) Sexual Relationships* 0.85-0.93; (vi) *Social Isolation* 0.90-0.96; (vii) *Coping* 0.84-0.91. Correlations between factors ranged from 0.64 (*Relationships (Family and Friends)* with *Coping*) to 0.96 (*General Mental Health* with *Self Worth*). This data provides strong support for 'convergent validity' (items loaded strongly by the theorized latent variable) and 'discriminant validity' (items loaded only by the theorized latent variable) of the final PsyQoL measurement model.

A Hierarchical Model of Psychological Quality of Life

It was postulated that the various components of psychological quality of life identified in the above analyses would be sufficiently interrelated to form a coherent structure whereby the individual components would be uniquely related to a single generative construct. This view was supported by the strong correlations between the components in the full measurement model. It was further evaluated by testing a three-level hierarchical model in which *Psychological Quality of Life* was causally related to *General Mental Health*, *Happiness*, *Self Worth*, *Relationships*, and *Coping*. The hypothesised *Relationships* factor was subsequently causally related to *Relationships (Social and Family)*, *Sexual Relationships*, and Social *Isolation*. The lowest-order latent variables were subsequently uniquely related to their defining items. Apart from the causal relationships specified above, no cross-loadings, correlations between latent variables (correlated disturbances) or correlations between item residuals were allowed.

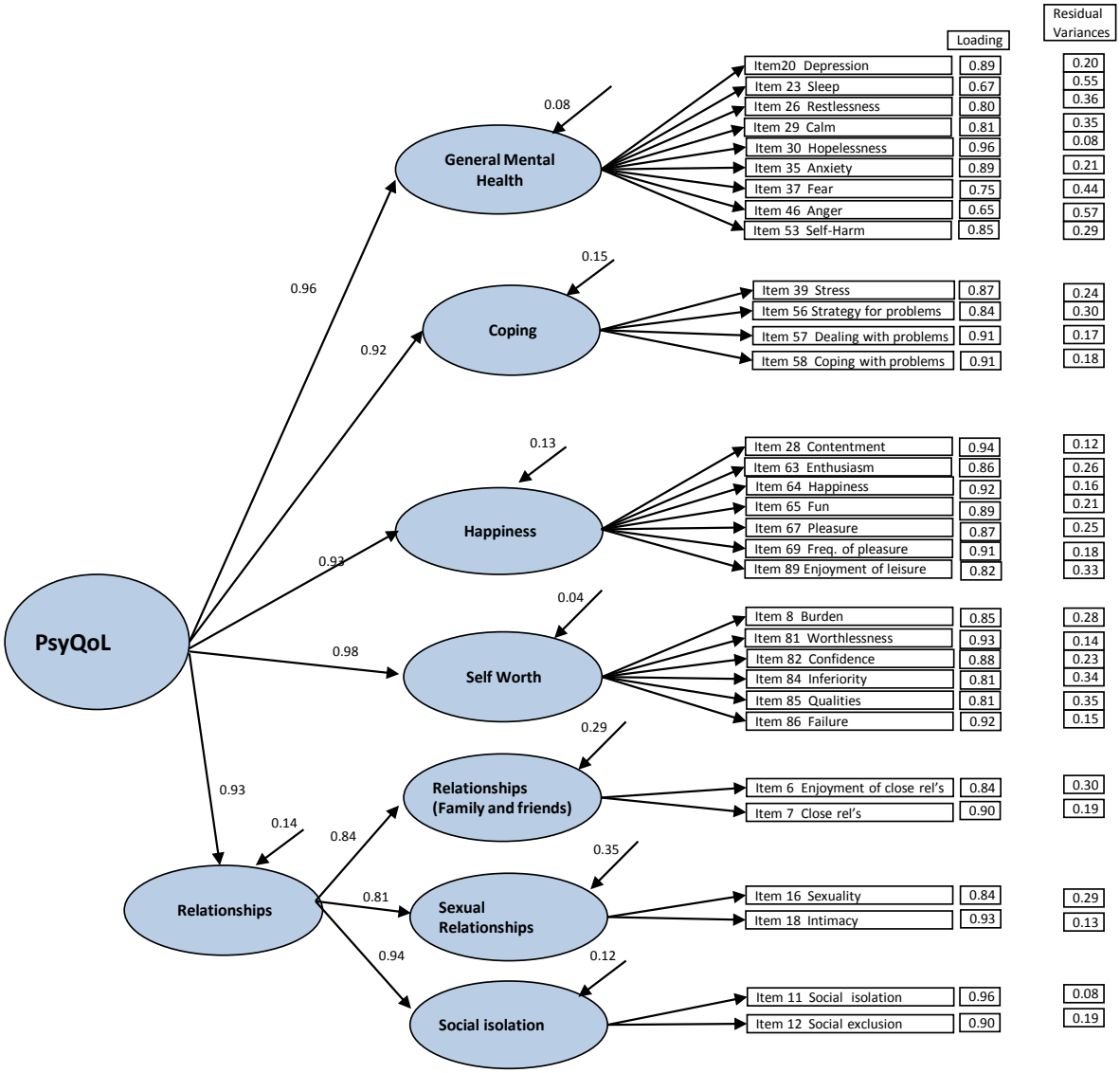
Considering the very restrictive nature of the hierarchical model the fit was satisfactory, (CFI = 0.981, TLI = 0.979, RMSEA = 0.075, WRMR = 1.499) and thus provided convincing evidence for

the coherence of a general *Psychological Quality of Life* construct – 'PsyQoL'. Additional features of the model were:

- Loadings of *Psychological Quality of Life* on the second-order constructs ranged from 0.92 to 0.98;
- Consequent residual variances of the second-order constructs although all statistically significant, were low, ranging from 0.04 *(Self Worth)* to 0.15 *(Coping)* and thus the proportions of variance explained in these constructs ranged from 96% to 85%;
- Loadings of the generic Relationships construct on the three specific relationships factors were 0.84 *(Relationships (Family and Friends)* 0.81 *(Sexual Relationships)* and 0.94 *(Loneliness)*;
- Residual variances of the three specific relationships factors were 71% *(Relationships (Family and Friends))*, 65% *(Sexual Relationships)* and 88% *(Loneliness)*;
- Loadings of the lower-order constructs on the specific items ranged from 0.65 (*General Mental Health* on Item 46) and 0.96 (*General Mental Health* on Item 30);
- Proportions of variance in the constituent items explained by the model (R-Squared) ranged from 43% (Item 46 - How often do you feel angry?) to 0.92 (Item 11 - How often do you feel socially isolated?).

The PsyQoL hierarchical model is shown in Figure 4.

Figure 4 The hierarchical PsyQoL Model (32 items)



Fit Statistics: CFI = 0.981; TLI = 0.979; RMSEA = 0.075; WRMR = 1.45

Notes:

Coefficients with unjoined arrows are the residual (unexplained) variances of the latent variables.

3.3 PsyQoL-Brief and AQoL-8D

PsyQoL-Brief

Following the development of the PsyQoL inventory a 'PsyQoL-Brief' version was constructed as a preliminary step towards integration with the AQoL-6D inventory. A sub-set of items from each PsyQoL dimension was selected by the research team for this purpose. It was aimed to develop a short scale of four items for each construct. The scales and items are summarised below. Item numbers relate to Table 4 and factor loadings are in brackets.

General Mental Health: Item 23 sleep (0.67); Item 35 anxiety (0.87); Item 46 anger (0.71); Item 53 self harm (0.83).

The one-factor congeneric model was an excellent fit: CFI = 1.000, TLI = 1.002, RMSEA = 0.000, WRMR = 0.045. Chi-square was also not significant (0.139, 2 d.f., pr = 0.9327). Internal consistency (Cronbach's alpha) = 0.80.

Happiness: Item 28 contentment (0.92); Item 63 enthusiasm (0.86); Item 64 happiness (0.92); Item 69 pleasure (0.91).

The fit of this model was not completely satisfactory (CFI = 0.999, TLI = 0.998, RMSEA = 0.090, WRMR = 0.400). There was, however, no obvious indication for model improvement. Cronbach's alpha = 0.92.

S*elf Worth*: Item 8 feeling a burden (0.84); Item 81 worthlessness (0.93); Item 82 confidence (0.84); Item 86 failure (0.93).

The one-factor congeneric model was an excellent fit (CFI = 1.000, TLI = 1.000, RMSEA = 0.000, WRMR = 0.138 – additionally Chi-square = 1.344, 2 d.f. pr = 0.5107). Cronbach's alpha = 0.91.

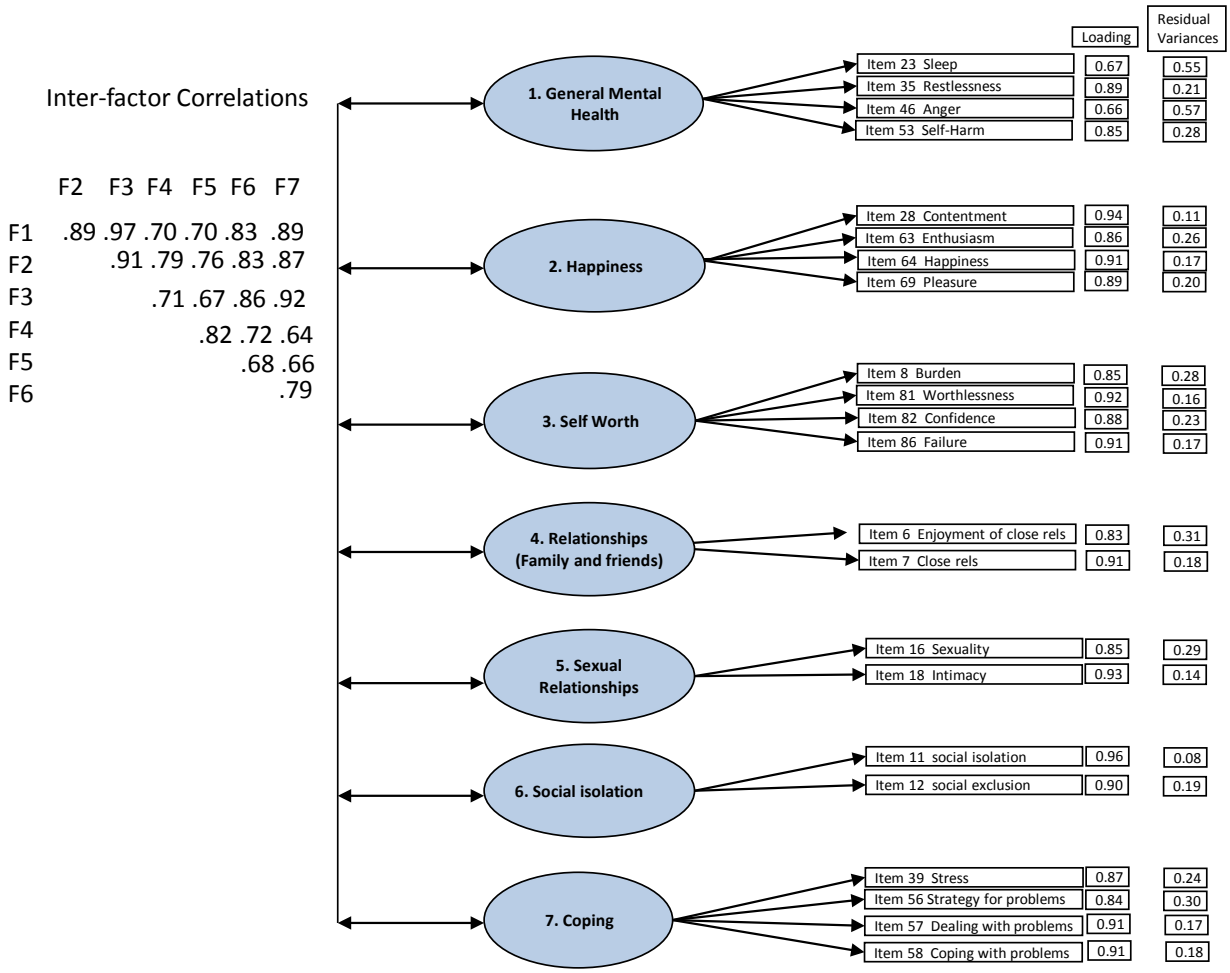
It was not possible to locate a model that adequately represented the three sub-domains of *Social and Family Relationships* with fewer than the six items selected for PsyQoL. These were therefore retained for PsyQoL-Brief. Similarly, the one-factor congeneric model for the four item *Coping* scale fitted the data very well, and were also retained.

A seven-factor measurement model incorporating the 'brief' item sets is shown in Figure 5. It fitted the data well (CFI = 0.992, TLI = 0.991 RMSEA = 0.058, WRMR = 0.919) i.e. all fit indices were within an acceptable range.

Additional features of the model included:

- Factor loadings ranged from 0.67 to 0.89 (*General Mental Health*), 0.86-0.94 (*Happiness*), 0.85-0.92 (*Self Worth*), 0.83-0.91 (*Relationships (Social and Family*)), 0.85-0.93 (*Sexual Relationships*), 0.90-0.96 (*Social Isolation*), and 0.84-0.91 (*Coping*);
- Correlations between factors ranged from 0.64 (*Relationships (Social and Family*) with *Coping)* to 0.97 (*General Mental Health* with *Self Worth*);
- Proportions of variance in the constituent items accounted for by the model ranged from 43% (Item 46 - anger) to 92% (Item 11 – isolation).

Figure 5 The PsyQoL-Brief Model (22 items)



Fit Statistics: CFI – 0.992; TLI = 0.991; RMSEA = 0.058; WRMR = 0.919

Notes:
Coefficients with unjoined arrows are the residual (unexplained) variances of the latent variables.

PsyQoL-Brief integrated with AQoL-6D

The possibility of integration of items with the AQoL-6D was examined for each of the PsyQoL dimensions separately and jointly.

General Mental Health: Four PsyQoL and four AQoL-6D items were combined: PsyQoL [23 – sleeping; 35 – anxiety; 46 – anger; 53 – self harm]; AQoL-6D [8 – despair; 9 – worry; 10 – sadness; 11 – tranquillity].

Model fit and internal consistency were good, CFI = 0.996, TLI = 0.995, RMSEA = 0.068, WRMR = 0.754; Cronbach's alpha = 0.92.

Happiness: There is not an analogous scale in AQoL-6D. Consequently the *Happiness* scale from PsyQoL-Brief was adopted (PsyQoL [28 – contentment; 63 – enthusiasm; 64 – happiness; and 69 – pleasure]).

Self Worth: Similarly, there is no equivalent scale in AQoL-6D and the revised set of four PsyQoL *Self Worth* items in PsyQoL-Brief were adopted (PsyQoL [8 – burden; 81 – worth; 82 – confidence; and 86 – failure]).

Social and Family Relationships: There are three items relevant to this construct in AQoL-6D: AQoL-6D items 5, 6, and 7. To achieve integration, the best fitting item from each of the three sub-sets of items in PsyQoL-Brief were selected: PsyQoL [7, 11 and 18]. The resulting six-item scale had very poor fit characteristics (CFI = 0.950, TLI = 0.917, RMSEA = 0.261, WRMR = 2.707). The addition of a number of correlations among the residuals was required to achieve an adequate fit, indicating considerable factorial complexity among this collection of items. The best model that could be located consisted of: PsyQoL 6 and 7 *Relationships* (*Family and Friends*); PsyQoL 16 and 18 (*Sexual Relationships*); AQoL-6D 7 *Health and Social Role*; PsyQoL 11 and 12 *Social Isolation*. This four-factor model had quite good fit (CFI = 0.996, TLI = 0.993, RMSEA = 0.066, WRMR = 0.626) and the sub-structure appears theoretically plausible. The internal consistency (Cronbach alpha) of the six items identified above plus AQoL-6D is 0.92.

Coping: There are three '*Coping*' items in AQoL-6D: items 12, 13 and 14. Combining these with the four coping in PsyQoL resulted in a single factor model which did not fit the data well (CFI = 0.993, TLI = 0.990, RMSEA = 0.122, WRMR = 1.136) Residual correlation indicated factorial complexity. One model that fitted the data quite well included two factors that separated the items from the contrasting inventories (AQoL-6D and PsyQoL). However, the 2 latent variables have a correlation of 0.94. As items were randomly distributed across the whole experimental questionnaire, and three random orders were used, it is unlikely that this can be considered an artefact of the source of the items. Consequently it was decided that the seven items be combined into a single scale. The resulting composite has considerable breadth, encompassing the ideas of handling stress, dealing and coping with problems, having the energy to do things you want to do, and feeling in control of life. The internal consistency of this composite scale is 0.93.

The integrated AQoL-6D/PsyQoL-Brief model described above would result in an inventory with 42 items. The first step in validating this inventory was to confirm the six-factor AQoL-6D model on the new data set. While some modifications were clearly possible a model with no cross-loadings or correlated residuals fitted the data reasonably well (CFI = 0.985, TLI = 0.982, RMSEA = 0.065, WRMR = 1.170).

The second step was to include the 22 recommended PsyQoL-Brief items to generate an eleven-factor model comprising: (a) *General Mental Health* (8 items) ; (b) *Social and Family Relations* (9 items) ; (c) *Coping* (7 items) *Happiness* (4 items); *Self Worth* (4 items); *Independent Living*, *Pain*, *Sensory Perception* (4, 3, and 3 items respectively). The fit of this eleven-factor model was reasonably good (CFI = 0.985, TLI = 0.983, RMSEA = 0.052, WRMR = 1.201) particularly considering that no cross-loadings or correlated errors were included. A 'part hierarchical' model in which the four social and family relations sub-factors were loaded by a single second-order factor that was then correlated with the other first-order factors also fitted the data reasonably well (CFI = 0.980, TLI = 0.978, RMSEA = 0.060, WRMR = 1.416). Additional details of this latter model include:

- Factor loadings on individual items ranged from 0.66 to 0.90 (*General Mental Health*), 0.87-0.94 (*Life Meaning and Satisfaction*), 0.86-0.92 (*Self Worth*), 0.82-0.91 (*Relationships - Social and Family*), 0.86-0.88 (*Sexual Relationships*), 0.90-0.96 (*Social Isolation*), 0.86-0.88 (*Health and Social Role*) 0.84-0.91 (*Coping*), 0.63-0.89 (*Independent Living*), 0.88-0.92 (*Pain*) and 0.39-0.83 (*Senses*);

- Loadings of the general *Social Relationships* factor on its constituent specific factors were 0.82 (*Relationships - Social and Family*), 0.81 (*Sexual Relationships*), 0.92 (*Social Isolation*) and 0.93 (*Health and Social Role*);
- Correlations between the second-order factors ranged from 0.54 (*Pain* with *Happiness*) to 0.95 (*Coping* with *Self Worth*);
- Proportions of variance in the constituent items accounted for by the model ranged from low values of 15% (AQoL-6D Item 18 – vision) and 30% (AQoL-6D Item 19 – hearing) to 92% (construction item bank Item 11 – isolation).

The hypothesis that scales in this integrated model would be related to a single generative construct (HRQoL) was tested with a three-level hierarchical model in which HRQoL was causally related to *General Mental Health*, *Happiness*, *Self Worth*, *Relationships - Coping, Independent Living, Pain and Senses*. The general *Relationships* factor - was causally related to *Relationships (Social and Family)*, *Sexual Relationships*, *Social Isolation* and *Health and Social Role* and HRQoL was causally related to general Relationships and the other 7 scales (see Figure 5). Apart from the causal relationships specified above, no cross-loadings, correlations between the residual variances of the latent variables (correlated disturbances) or correlations between item residuals were initially allowed.

Considering the tight restrictions imposed, the model fitted the data reasonably well (CFI = 0.973, TLI = 0.971, RMSEA = 0.068, WRMR = 1.669). Additional features of the model were:

- Loadings of Health-related Quality of Life on the second-order constructs ranged from 0.64 (*Pain*) to 0.97 (*Self Worth*);
- Consequent residual variances of the second-order constructs were, with two exceptions, low, ranging from 0.05 (*Self Worth*) to 0.53 (*Senses*) and 0.59 *(Pain*) and thus the proportions of variance explained in these constructs ranged from 95% (*Self Worth*) to 41% (*Pain*) and 47% (*Senses)*;
- Loadings of the general *Relationships* - construct on the four specific relationships factors were 0.82 (*Relationships - Family and Friends*) 0.81 (*Sexual Relationships*), 0.92 (*Social Isolation*) and 0.93 (*Health and Social Role*);
- Proportions of variance explained in the four specific relationships factors were 67% (*Relationships - Family and Friends*), 65% (*Sexual Relationships*), 84% (*Social Isolation*) and 87% (*Health and Social Role*);
- Loadings of the lower-order constructs on the specific items ranged from 0.39 (*Senses* on AQoL Item 18) and 0.56 (*Senses* on AQoL Item 19) to 0.96 (Social Isolation on Item 11);
- Proportions of variance in the constituent items explained by the model (R^2) ranged from low values of 15% (AQoL-6D Item 18 – vision and 30% (AQoL-6D Item 19 – hearing) to 0.92 (construction item bank Item 11 – isolation).

While the hierarchical model fitted the data well, the results suggested that the two AQoL 2 dimensions that tap physical symptoms, *Pain* and the 'vision' and 'hearing' components of the *Senses* construct, may not be particularly strongly related to the more social and psychological components of a conception of health-related quality of life. In contrast, the AQoL-6D dimension connoting Independent Living appeared to be quite strongly related to the more social and psychological constructs.

With such a restricted model a number of modifications that would potentially improve model fit were possible. For example fitting two additional parameters with the clearly highest modification indices improved the fit to CFI = 0.978, TLI = 0.977, RMSEA = 0.061, WRMR = 1.501. The WRMR remained above the recommended cut-off however.

AQoL-8D and Super Dimensions

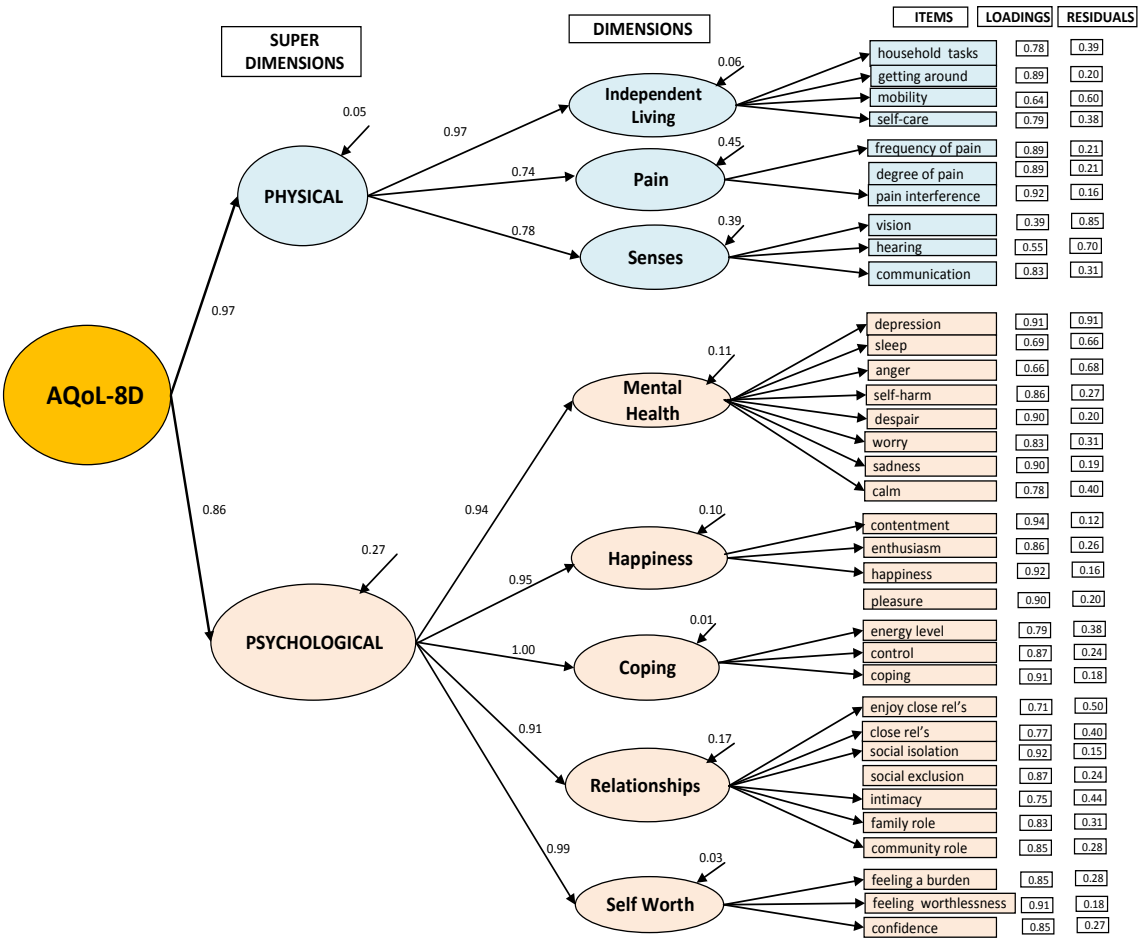
The final stage of the item selection sought to optimise the brevity of the instrument. The analysis was based primarily upon substantive arguments: which items from the combined AQoL-6D/PsyQoL-Brief could be dropped without compromising the theoretical concept and without excessive cost in terms of the deterioration in the statistical fit. The super-imposition of the items of AQoL-6D upon those generated for PsyQoL-Brief resulted in a number of items which were judged on substantive grounds, to be at least partially redundant. With the inclusion of 'despair', 'worry' and 'sadness' from AQoL-6D, 'anxiety' was judged unnecessary. Similarly the AQoL-6D items for *Coping* – 'energy' and 'control' displaced three items from PsyQoL-Brief. 'Failure' was omitted from *Self Worth* owing to its correlation with 'worthlessness' and 'sexuality' was omitted from *Relationships* because of its detrimental effect upon response rates. 'Depression' – item 20 – was added back into the instrument since, without it, the instrument would lack face validity.

An eight-factor 'congeneric model' was fitted to the data for the resulting 35 items. Each item was uniquely associated with one construct only and no cross-loadings or correlations among residuals were allowed. The model fitted the data reasonably well (CFI = 0.976, TLI = 0.973, RMSEA = 0.071, WRMR = 1.513) although the WRMR is above the recommended range. Other features of the model illustrated in Figure 6 include:

- Factor loadings on individual items ranged from 0.67 to 0.91 (*General Mental Health*), 0.86-0.94 (*Happiness*), 0.85-0.91 (*Self Worth*), 0.71-0.92 (*Relationships - General*), 0.84-0.91 (*Coping*), 0.64-0.89 (*Independent Living*), 0.89-0.92 (*Pain*) and 0.39-0.83 (*Senses*);
- Correlations between the factors ranged from 0.54 (*Pain* with *Happiness*) to 0.95 (*Coping* with *Self Worth*);
- Proportions of variance in the constituent items accounted for by the model ranged from low values of 16% (AQoL-6D Item 18 - Thinking about your vision (using your glasses or contact lenses if needed): I have excellent sight … etc.) and 30% (Thinking about your hearing (using your hearing aid if needed): I have excellent hearing … etc.) to 85% (Item 11 - How often do you feel socially isolated?).

The 8 factor model was also fitted to a 2 factor model consisting of two 'super dimensions' labelled 'Physical' and 'Psychological'. With only two second order factors loaded by AQoL identification is not possible. However the correlation between the two factors is 0.83 which permits construction of the model shown in Figure 6. In this, the construct explains 'Physical' 55%, 61% and 94% of the variances of the *Pain*, *Senses* and *Independent Living* dimensions. The 'Psychological' construct explains between 83% and 99% of the variances of the remaining dimensions.

Figure 6 The AQoL-8D model



Fit Statistics: CFI = 0.974; TLI = 0.972; RMSEA = 0.073; WRMR = 1.64

Notes:

1. Unstandardised loadings of AQoL-8D on Physical and Psychological were constrained to be equal identification.
2. Coefficients with unjoined arrows are the residual (unexplained) variances of the latent variables.

4. Properties of AQoL-8D

The interval consistency of AQoL-8D and the 8 dimensions was tested using Cronbach alpha. Results reported in Table 5 are satisfactory except for 'Senses' where items were retained for substantive reasons despite the poor fit.

The properties of the final instrument were tested two additional ways: (i) by comparing scores for patients and the general population; and (ii) by comparing individuals where AQoL-8D produced a greater numerical score than AQoL-6D with individuals where AQoL-6D exceeded AQoL-8D. Item

scores were summed to obtain dimension scores, each of which was then transformed to a 0-1 scale. Total scores were obtained by summing dimension scores and similarly transforming them to a 0-1 scale.

Table 5 Cronbach alpha for the AQoL-8D dimensions

Dimension	α	Dimension	α
Independent living	0.81	Relationships	0.90
Life satisfaction	0.92	Self worth	0.87
Mental health	0.92	Pain	0.85
Coping	0.85	Senses	0.54
Physical super dimension	0.84	Psycho-social super dimension	0.96
AQoL-8D	0.97		

The prior expectations, for the first test were (1.1) that patients would obtain lower AQoL-8D overall scores and particularly for psycho-social dimensions; and (1.2) that neither the overall instrument nor sub-scales would have significant floor or ceiling effects. If AQoL-6D and AQoL-8D were equally sensitive to psycho-social factors (ie if the general items of AQoL-6D fully captured this content and AQoL-8D failed to measure extra psycho-social content) then for the subset of patients where AQoL-8D < AQoL-6D the average scores of physical and psycho social dimensions would be similar with the difference in instrument scores being attributable to random factors. If, however AQoL-8D is more sensitive to psycho social factors then: (2) these dimensions will have lower scores and physical dimensions higher scores (since scales on both instruments were standardised).

Figure 7 and Table 6 describe the AQoL-8D scores. By the standards of HR-MAU instruments the frequency distributions in Figure 7 are very satisfactory, approximating a normal distribution. Figure 7a indicates the absence of a ceiling effect and Figure 7b a low floor value. This is confirmed in Table 6 where 0.4% and 0.0% of patients and the public respectively attain the maximum possible score. A similar observation is true for the majority of dimensions.

Mean patient and public values from Table 6 are plotted in Figure 8. Differences between the two are statistically significant for every dimension and for AQoL-8D. The magnitude of the difference is largest for the mental health dimensions and particularly self worth. It is smallest for senses and pain.

The relationship between AQoL-6D and 8D is shown in Figures 9a and 9b. The Pearson correlation for patients and public are 0.965 and 0.936 respectively, reflecting the close but imperfect association. Differences between the two instruments are shown in the frequency distributions, Figure 10a and 10b. Both are negatively skewed reflecting the (statistically insignificant) difference of 0.02 between instrument scores for public patients. Both patients and public were separated into two groups: those where AQoL-8D > AQoL-6D, and those where AQoL-6D > AQoL-8D. Average AQoL-8D and dimension scores for the two groups are shown in Figures 11a and 11b and Table 7. The differences between instruments exactly parallel the differences between the public and patient groups analysed earlier. The difference between the populations is statistically significant for every dimension and for the overall AQoL-8D score. When AQoL-8D scores are lower than AQoL-6D all of the mental health dimensions and especially self worth have smaller mean values and by a large magnitude. Mean values for pain and senses are larger, again reflecting the standardisation of scores. The result confirms that differences between the instruments are generated by the greater sensitivity of AQoL-8D in the domains of mental health.

Table 6 Descriptive statistics AQoL-8D

Psychometric score and respondent		N	Mean	Std err	95% confidence interval for mean		Min	Max	% score = Max (ceiling effect)
					Lower bound	Upper bound			
AQoL-6D	Public	197	.82	.006	.81	.84	.51	.97	0.0
	Patient	512	.61	.007	.59	.62	.20	1.00	0.4
	Total	709	.67	.007	.65	.68	.20	1.00	0.3
AQoL-8D	Public	197	.80	.006	.79	.81	.50	.97	0.0
	Patient	509	.57	.008	.55	.58	.20	1.00	0.4
	Total	706	.63	.007	.62	.65	.20	1.00	0.3
Independent Living	Public	197	.91	.008	.89	.92	.28	1.00	34.0
	Patient	509	.69	.008	.68	.71	.22	1.00	6.5
	Total	706	.75	.007	.74	.77	.22	1.00	14.2
Happiness	Public	197	.72	.009	.70	.73	.38	1.00	1.0
	Patient	509	.47	.010	.45	.49	.00	1.00	2.4
	Total	706	.54	.009	.52	.56	.00	1.00	2.0
Mental health	Public	197	.75	.009	.73	.77	.27	.97	0.0
	Patient	509	.51	.010	.49	.53	.03	1.00	1.2
	Total	706	.58	.009	.56	.59	.03	1.00	0.8
Coping	Public	197	.74	.010	.72	.75	.17	1.00	4.1
	Patient	509	.47	.011	.45	.49	.00	1.00	2.6
	Total	706	.55	.009	.53	.56	.00	1.00	3.0
Relationships	Public	197	.84	.008	.82	.85	.48	1.00	10.7
	Patient	509	.57	.009	.55	.59	.00	1.00	1.2
	Total	706	.64	.008	.63	.66	.00	1.00	3.8
Self worth	Public	197	.81	.009	.79	.83	.42	1.00	8.1
	Patient	509	.49	.012	.47	.51	.00	1.00	3.1
	Total	706	.58	.010	.56	.60	.00	1.00	4.5
Pain	Public	197	.81	.013	.78	.84	.10	1.00	28.4
	Patient	509	.63	.012	.61	.66	.00	1.00	18.1
	Total	706	.68	.010	.66	.70	.00	1.00	21.0
Senses	Public	197	.85	.007	.84	.87	.54	1.00	14.2
	Patient	509	.77	.006	.75	.78	.31	1.00	9.6
	Total	706	.79	.005	.78	.80	.31	1.00	10.9
Physical Super Dimension	Public	197	.87	.007	.85	.88	.41	1.00	4.1
	Patient	509	.70	.007	.69	.71	.24	1.00	1.6
	Total	706	.75	.006	.74	.76	.24	1.00	2.3
Mental Super Dimension	Public	197	.77	.007	.76	.79	.43	.97	0.0
	Patient	509	.51	.009	.50	.53	.06	1.00	0.6
	Total	706	.59	.008	.57	.60	.06	1.00	0.4

Data source: AQoL-8D construction database (Public 197; Patient 516)

Table 7 AQoL-8D minus AQoL-6D: analysis of differences

Psychometric score and Population group		N	Mean	Std error	95% Confidence Interval for Mean	
					LB	UB
AQoL-8D	Pop 1 (8D>6D)	110	.69	.016	.655	.720
	Pop 2 (6D>8D)	526	.62	.008	.601	.634
	Pop 3 (8D=6D)	70	.66	.023	.614	.707
	Total	706	.63	.007	.619	.647
Happiness	Pop 1 (8D>6D)	110	.68	.019	.642	.716
	Pop 2 (6D>8D)	526	.50	.010	.481	.521
	Pop 3 (8D=6D)	70	.60	.028	.547	.659
	Total	706	.54	.009	.521	.556
Mental Health	Pop 1 (8D>6D)	110	.67	.018	.633	.705
	Pop 2 (6D>8D)	526	.55	.010	.532	.572
	Pop 3 (8D=6D)	70	.63	.026	.572	.678
	Total	706	.58	.009	.561	.594
Coping	Pop 1 (8D>6D)	110	.64	.021	.598	.682
	Pop 2 (6D>8D)	526	.52	.011	.503	.544
	Pop 3 (8D=6D)	70	.56	.030	.500	.619
	Total	706	.55	.009	.527	.563
Relationships	Pop 1 (8D>6D)	110	.74	.018	.707	.780
	Pop 2 (6D>8D)	526	.62	.010	.599	.638
	Pop 3 (8D=6D)	70	.69	.024	.640	.736
	Total	706	.64	.008	.629	.661
Self Worth	Pop 1 (8D>6D)	110	.73	.021	.691	.775
	Pop 2 (6D>8D)	526	.54	.012	.512	.559
	Pop 3 (8D=6D)	70	.68	.029	.620	.734
	Total	706	.58	.010	.560	.600
Pain	Pop 1 (8D>6D)	110	.56	.025	.513	.612
	Pop 2 (6D>8D)	526	.71	.011	.691	.734
	Pop 3 (8D=6D)	70	.65	.032	.583	.709
	Total	706	.68	.010	.664	.702
Senses	Pop 1 (8D>6D)	110	.75	.014	.718	.775
	Pop 2 (6D>8D)	526	.80	.006	.789	.812
	Pop 3 (8D=6D)	70	.78	.016	.748	.811
	Total	706	.79	.005	.780	.800
Physical Super Dimension score	Pop 1 (8D>6D)	110	.67	.016	.638	.701
	Pop 2 (6D>8D)	526	.77	.007	.754	.781
	Pop 3 (8D=6D)	70	.72	.020	.677	.759
	Total	706	.75	.006	.735	.759
Mental Health Super Dimension score	Pop 1 (8D>6D)	110	.69	.017	.660	.729
	Pop 2 (6D>8D)	526	.56	.009	.537	.575
	Pop 3 (8D=6D)	70	.64	.025	.587	.686
	Total	706	.59	.008	.570	.602

Figure 7 Frequency Distribution of AQoL-8D Psychometric Score

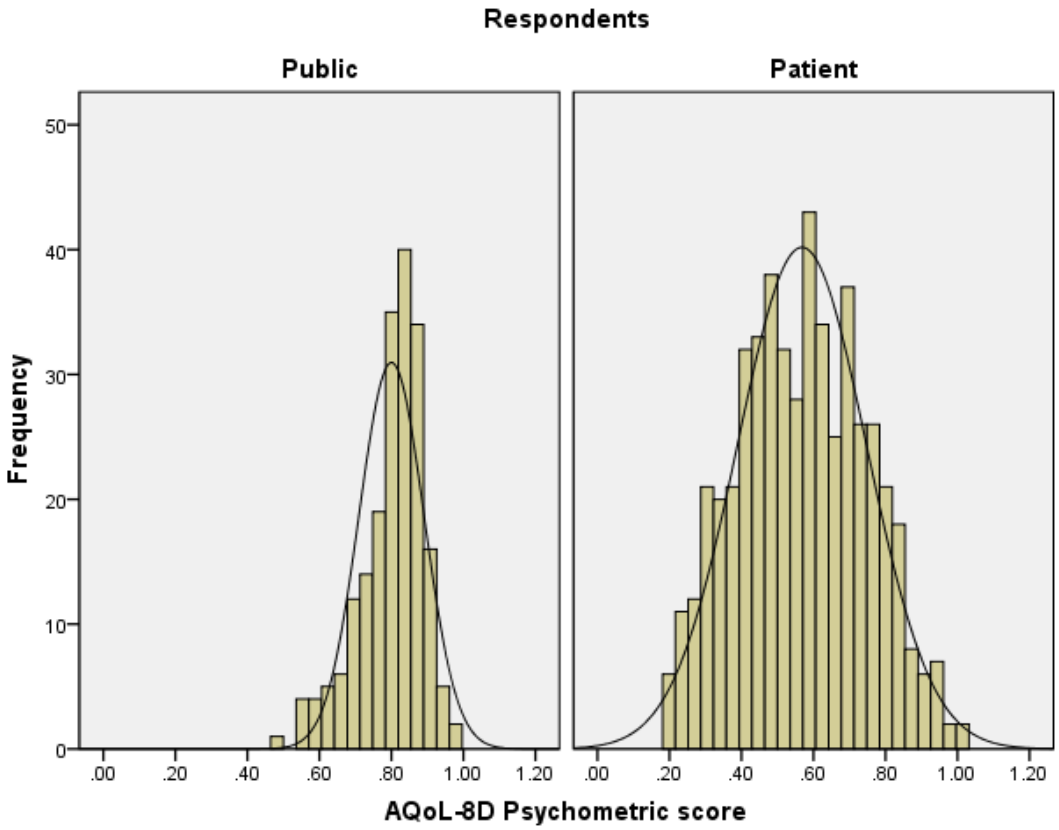


Figure 8 AQoL-8D and Dimensions Scores by Respondent Group

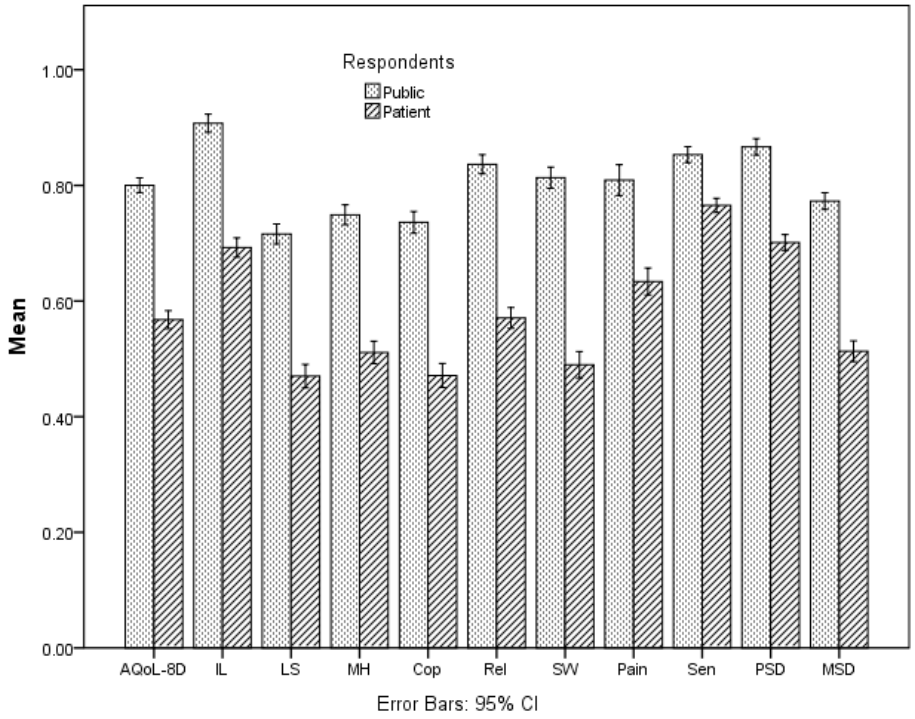


Figure 9a Comparison of AQoL-6D and AQoL-8D (public)

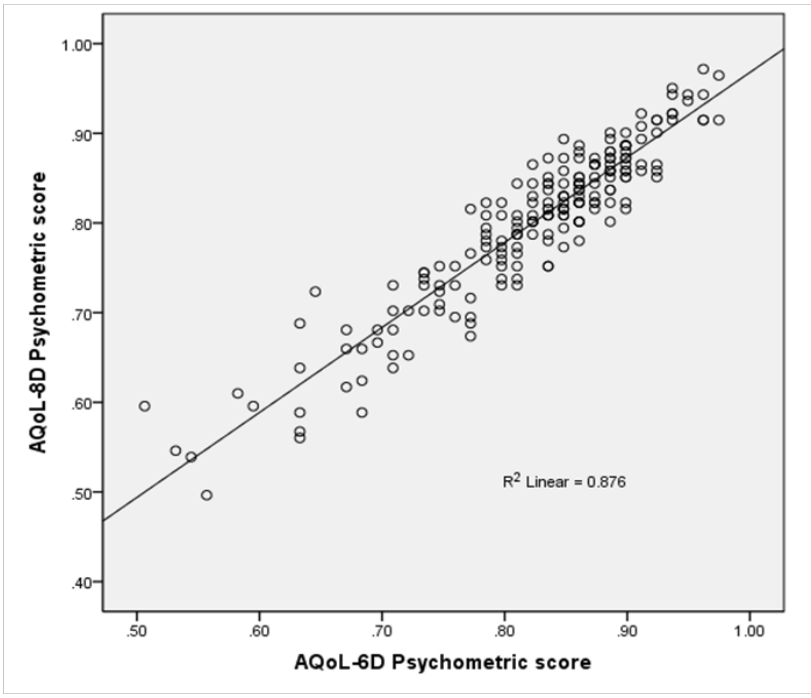


Figure 9b Comparison of AQoL-6D and AQoL-8D (patient)

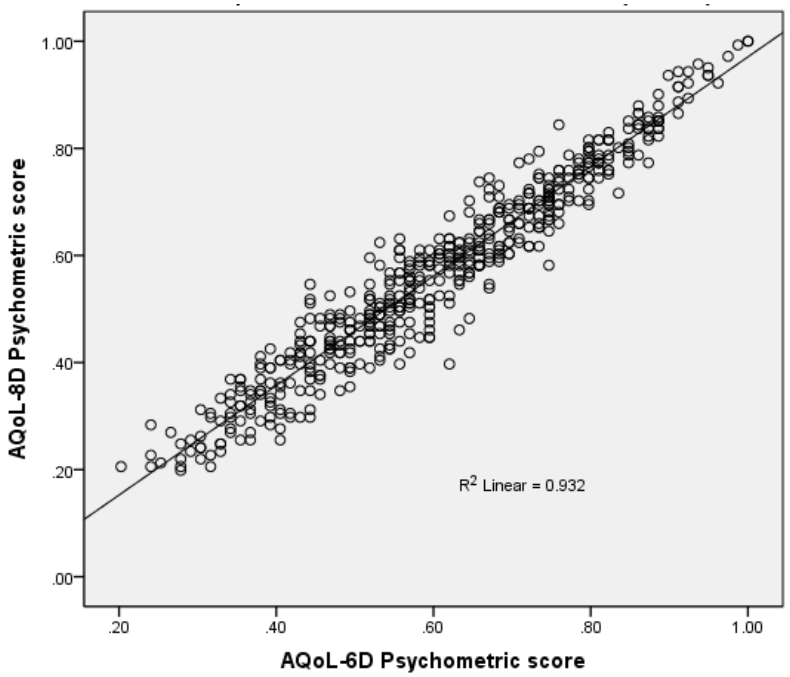


Figure 10a Frequency distribution between instrumental scores (public)

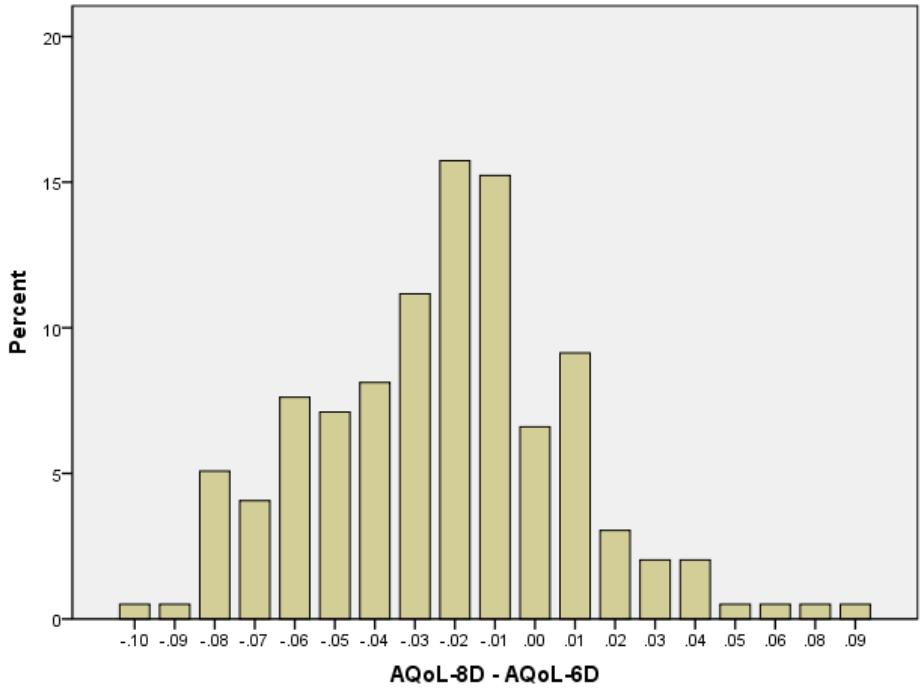


Figure 10b Frequency distribution between instrumental scores (patient)

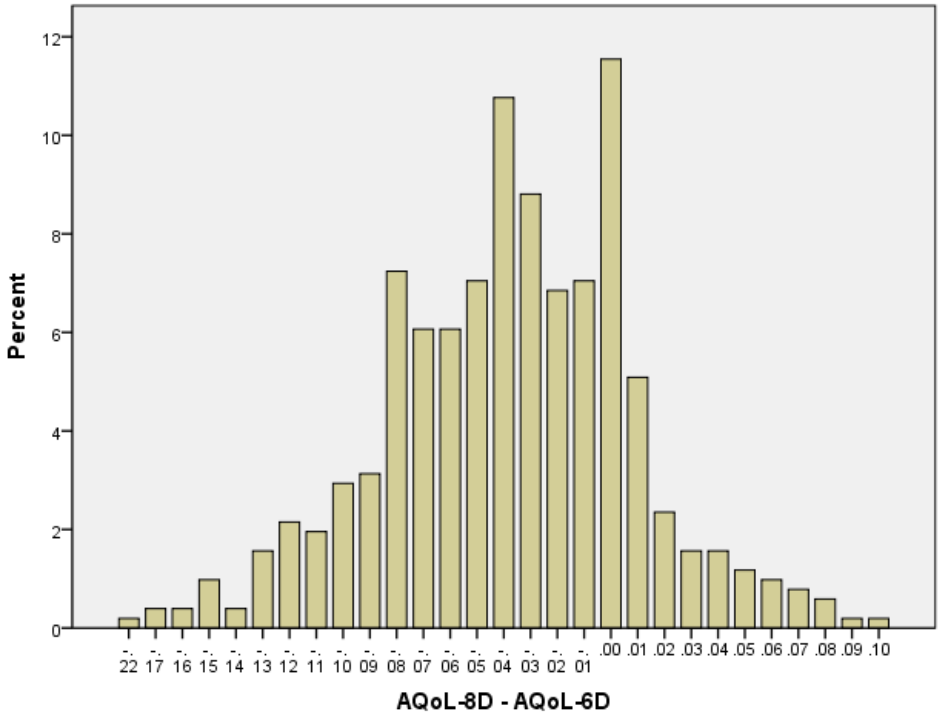


Figure 11a AQoL-8D: Mean scores when AQoL-8D >< AQoL-6D (general population)

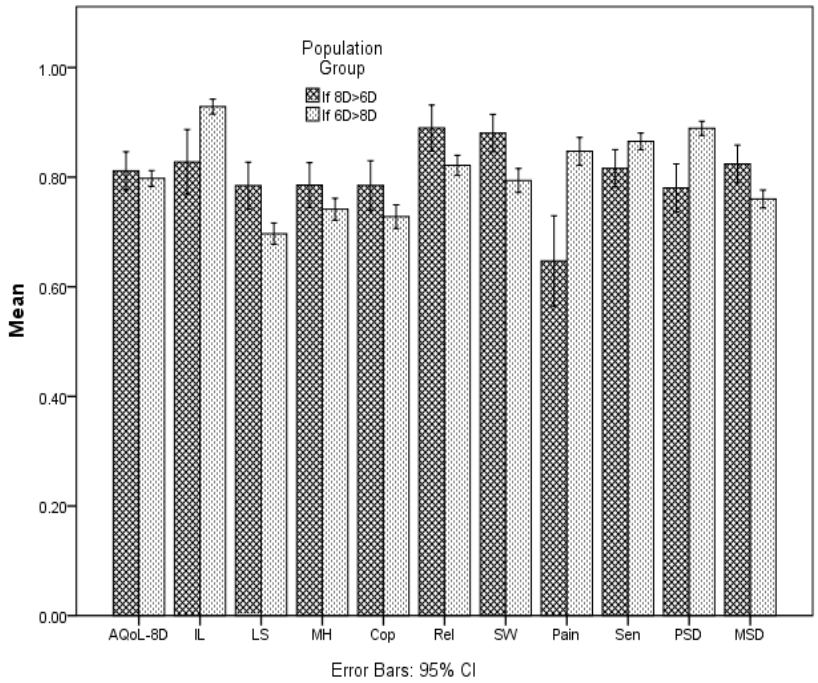
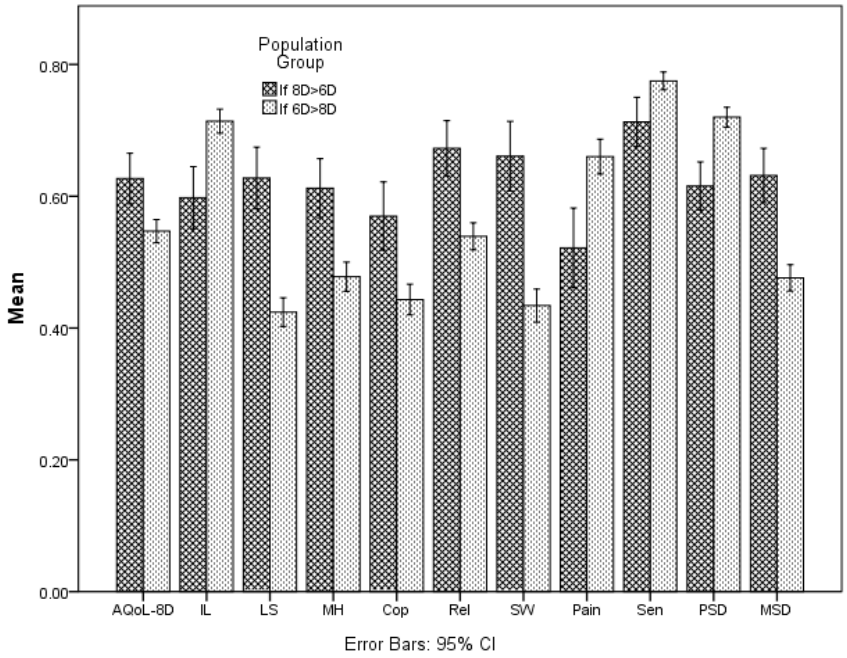


Figure 11b Mean scores when AQoL-8D >< AQoL-6D (patients)



5 Discussion and Conclusions

According to psychometric theory scales should be constructed from the interplay between theoretical concerns (expert knowledge of a subject) and the statistical relationships between indicated items. The AQoL instruments were the first of the generic HRQoL instruments designed specifically for economic evaluation to seek content validity in this way. The emphasis in the health economics literature has been almost exclusively upon the properties of the weights to be applied to items and the combination models. Content validity has been assumed.

The AQoL instruments commenced with a relatively small inventory, reflecting in part, the beliefs that generic QoL instruments for evaluating all health related interventions can be brief and must be brief (or clinicians will not use them). These beliefs are unsubstantiated (and it is questionable whether clinicians would use an invalid brief instrument if they were not told that the instrument was acceptable). AQoL-6D was a response to content insensitivity in the social and mental domains and the present instrument was an extension of this work.

By the standards of MAUI the AQoL-8D is large and takes an average of 5.4 minutes to complete on-line. This is a reflection of the breadth of the construct and the added complexity arising from the inclusion of dimensions which have a limited psychometric relationship but are included in the common concept of health and are relevant for health policy.

The present paper demonstrates how these constraints may be reconciled with rigorous psychometric analysis. The final instrument is a compromise driven by substantive concerns and the pressure for instrument brevity. However results demonstrate that the AQoL-8D has good psychometric properties: that it describes an underlying latent variable which explains a significant part of the variation in dimensions and item scores. From its construction, the latent variable is a measure of HRQoL conceptualised as handicap.

The AQoL instruments add to the growing number of generic instruments available for measuring HRQoL which do not produce identical scores. For many purposes, the existence of multiple and numerically inconsistent instruments is unproblematical. These purposes include classification and ordinal ranking of health states. Differences are also unproblematical when each instrument has specialised use and comparison is not necessary. Instruments developed as MAUI for economic evaluation differ. Their *raison d'etre* is the comparison of dissimilar interventions and the ranking of projects. Differences in instrument scores alter the apparent benefits of one project relative to another.

Differences can arise because of the use of different utility weights or because of differences in the content of instruments. The former issue has been discussed at length by economists, but the second has been almost entirely ignored. However the comparisons of content which have been undertaken – including the comparison of AQoL-6D and AQoL-8D above – reveal significant differences. Until this problem has been successfully resolved the use of QALYs in economic evaluation will introduce a random element into the ranking of possible projects. As a first step the methods developed by psychometricians must be adopted more widely to achieve content validity. The present paper is one example of how this may be achieved.

References

Böckerman, P., E. Johansson, et al. (2011). "Do established health-related quality-of-life measures adequately capture the impact of chronic conditions on subjective well-being?" Health Policy **100**: 91-95.

Brazier, J., J. Ratcliffe, et al. (2007). Measuring and Valuing Health Benefits for Economic Evaluation. Oxford, Oxford University Press.

Brazier, J. and J. Roberts (2004). "The estimation of a preference-based measure of health from the SF-12." Medical Care **42**(9): 851-859.

Brazier, J., J. Roberts, et al. (2002). "The estimation of a preference-based measure of health from the SF-36." Journal of Health Economics **21**: 271-292.

Brooks, R., R. Rabin, et al. (2005). The Measurement and Valuation of Health Status using EQ-5D: A European Perspective, Evidence from the EuroQoL BIOMED Research Programme. Dordrecht, Kluwer Academic Publishers.

Browne, M. W. (2001). "An overview of analytic rotation in Exploratory Factor Analysis." Multivariate Behavioral Research **36**(1): 111-150.

Browne, M. W. and R. Cudeck (1993). Alternative ways of assessing model fit. Testing Structural Equation Models. B. K and L. J. Newbury Park, Sage.

Browne, M. W., R. Cudeck, et al. (1998). CEFA: Comprehensive Exploratory Factor Analysis. Columbus, The Ohio State University, http://faculty.psy.ohio-state.edu/browne/software.php/ (Retrieved March 21, 2011).

Browne, M. W., R. Cudeck, et al. (2008). CEFA: Comprehensive Exploratory Factor Analysis, Version 3.03 (Computer software manual), http://faculty.psy.ohio-state.edu/browne/ (Accessed April 4, 2011).

Dolan, P., C. Gudex, et al. (1995). A social tariff for EuroQoL: Results from a UK general population survey, Discussion Paper No 138. York, Centre for Health Economics, University of York.

Fayers, P. M. and D. Machin (2000). Quality of Life: Assessment, Analysis and Interpretation. Chichester, John Wiley & Sons Ltd.

Feeny, D., W. Furlong, et al. (2002). "Multi attribute and single attribute utility functions for the Health Utilities Index Mark 3 System." Medical Care **40**(2): 113-128.

Fryback, D. G., M. Palta, et al. (2010). "Comparison of 5 health related quality of life indexes using item response theory analysis." Medical Decision Making **30**(1): 5-15.

Hancock, G. R. and R. O. Mueller (2006). Structural Equation Modelling: A Second Course. Greenwich CT, Information Age Publishing Inc.

Hawthorne, G. (2009). "Assessing utility where short measures are required: development of the short Assessment of Quality of Life 8 (AQoL 8) instrument." Value in Health **12**(6): 948-957.

Hawthorne, G., J. Richardson, et al. (2001). "A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments." Annals of Medicine **33**: 358-370.

Hawthorne, G., J. Richardson, et al. (1999). "The Assessment of Quality of Life (AQoL) instrument: A psychometric measure of health related quality of life." Quality of Life Research **8**: 209-224.

Hawthorne, G., J. Richardson, et al. (1999). "The Australian Quality of Life (AQoL) instrument: a psychometric measure of health related quality of life." Quality of Life Research **8**: 209-224.

Jöreskog, K. G. (2001). Analysis of Ordinal Variables Part 2: Cross-Sectional Data, Scientific Software International**:** http://www.ssicentral.com.

Kaplan, R. and J. P. Anderson (1996). The General Health Policy Model: An Integrated Approach. Quality of Life and Pharmacoeconomics in Clinical Trials. B. Spilker. Philadelphia, Lippinocott-Raven Publishers**:** 309-322.

Khan, M. A. and J. Richardson (2011). A comparison of 7 instruments in a small, general population, Research Paper 60. Melbourne, Centre for Health Economics, Monash University.

McDonald, R. P. (2005). "Semiconfirmatory Factor Analysis: The Example of Anxiety and Depression." Structural Equation Modeling: A Multidisciplinary Journal **12**(1): 163-172.

Mulaik, S. A. and R. E. Millsap (2000). "Doing the Four-Step Right." Structural Equation Modeling: A Multidisciplinary Journal **7**(1): 36-73.

O'Connor, R. (2004). Measuring Quality of Life in Health. Edinburgh, Churchill Livingstone.

Richardson, J., N. A. Day, et al. (2004). "Measurement of the quality of life for economic evaluation and the Assessment of Quality of Life (AQoL) Mark 2 Instrument." Australian Economic Review **37**(1): 62-88.

Richardson, J., J. McKie, et al. (2011). Multi Attribute Utility Instruments and their use, Research Paper 63, Centre for Health Economics, Monash University.

Sintonen, H. and M. Pekurinen (1989). "A generic 15 dimensional measure of health-related quality of life (15D)." Journal of Social Medicine **26**: 85-96.

Streiner, D. L. and G. R. Norman (2003). Health Measurement Scales: A practical guide to their development and use. Oxford, Oxford University Press.

Torrance, G., D. Feeny, et al. (1996). "Multiattribute utility function for a comprehensive health status classification system: Health Utilities Index Mark 2." Medical Care **34**(7): 702-722.

Yu, C.-Y. (2002). Evaluating Cut-off Criteria of Model Fit Indices for Latent Variable Models with Binary and Continuous Outcomes, Dissertation for PhD. PhD, University of California.