Deakin Research Online

Deakin University's institutional research repository

This is the published version (version of record) of:

Parsian, Nasrin and Dunning, Trisha 2009-04, Developing and validating a questionnaire to measure spirituality : a psychometric process, *Global journal of health science*, vol. 1, no. 1, pp. 2-11.

.

Available from Deakin Research Online:

http://hdl.handle.net/10536/DRO/DU:30019516

Developing and Validating a Questionnaire to Measure Spirituality: A Psychometric Process

Nasrin Parsian (corresponding author)

School of Nursing, Deakin University

Burwood Highway, Burwood 3125 Victoria, Australia

E-mail: npars@deakin.edu.au

Trisha Dunning AM

Deakin University and Barwon Health

Geelong, Victoria Australia

E-mail: tridun@BarwonHealth.org.au

Abstract

The purpose of the paper is to describe the processes undertaken to evaluate the psychometric properties of a questionnaire developed to measure spirituality and examine the relationship between spirituality and coping in young adults with diabetes.

The specific validation processes used were: content and face validity, construct validity using factor analysis, reliability and internal consistency using test-retest reliability and Cronbach's alpha correlation coefficient.

The exploratory factor analysis revealed four factors: self-awareness, the importance of spiritual beliefs, spiritual practices, and spiritual needs. The items on the Spirituality Questionnaire (SQ) revealed factor loading ≥ 0.5. Reliability processes indicated that the SQ is reliable: Cronbach's alpha 0.94 for the global SQ and between 0.80-0.91 for the four subscales. Test-retest statistic examination revealed stability of the responses at two time points 10 weeks apart.

The final questionnaire consists of 29 items and the psychometrics indicated that it is valid and reliable.

Keywords: Inner self, Spirituality, Reliability, Validity

1. Background

Spirituality is increasingly being recognised as an important aspect of the health and wellbeing of people with chronic health conditions. Spirituality gives meaning to people's lives and may be an important coping resource that enables people with chronic conditions to manage their condition (Cronbach & Shavelson 2004; Tse, Lloyd, Petchkovsky & Manaia 2005). In addition spirituality is central to finding meaning, comfort and inner peace, which helps people transcend their condition and incorporate it into their self-concept (transformation). However, several barriers prevent spirituality from being incorporated into health care. For example, there is no consensus definition of 'spirituality' (McSherry & Draper 1998). The difficulty in defining spirituality is partly due to the fact that it is complex, highly subjective, and difficult to measure (Coyle 2002).

Currently, most validated spirituality tools concentrate on religion or higher beings and may only apply to religious people or those whose spirituality encompasses religion (Tuck, McCain & Elswick 2001). While religion is an aspect of spirituality for many people, but it is not synonymous with spirituality. Rather spirituality involves humans' search for meaning in life while religion usually involves rituals and practices and a higher power or 'God' (Tanyi 2002).

The current paper reports the processes used to develop and validate a spirituality questionnaire that focuses on the concepts of inner-self, meaning in life and connectedness to be used by young people with type 1 diabetes, to test the hypothesis that there is a relationship between spirituality and coping in young adults with diabetes. For the purpose of the study spirituality was defined as a concept encompassing finding meaning in life, self-actualisation and connection with inner self and the universal whole.

2. Method

The methods used to validate the SQ included:

− Translational validity: content validity and face validity.

− Construct validity: factor analysis.

− Reliability tests: internal consistency (Cronbach's alpha) and test-retest.

A flow chart depicting the processes used to examine the validity of the SQ is presented in Figure 1.

The draft spirituality questionnaire (SQ) was derived from the relevant literature and four existing 'spirituality' tools:

▪ The Spirituality Scale: The internal consistency of subscales ranged from 0.59 to 0.97 (Delaney 2005).

▪ The Spiritual Meaning Scale (SMS): Cronbach's alphas of the subscales ranged from 0.60 to 0.62 (Mascaro, Rosen & Morey 2004).

▪ Daily Spiritual Experiences Scale: Cronbach's alpha correlation coefficient for the global scale was 0.90 (Underwood, Institute & Teresi 2002).

▪ A survey used in a national study in the Higher Education Research Institute by the University of California to explore students' search for meaning and purpose. The internal consistency of the subscales ranged between 0.75 and 0.97. All of these scales were valid but focused on religion or higher being as a measure of spirituality and did not fit the definition of spirituality developed for the current study. In addition, they may not be relevant to non- religious people.

The initial draft of the SQ contained 35 items in seven sections:

▪ Importance of spiritual beliefs.

▪ Self-awareness.

▪ Environmental awareness.

▪ Relationships.

▪ Spiritual needs.

▪ Spiritual experiences.

▪ Open-ended questions.

2.1 Translational validity

2.1.1 Content validity

Content validity was undertaken to ascertain whether the content of the questionnaire was appropriate and relevant to the study purpose. Content validity indicates the content reflects a complete range of the attributes under study and is usually undertaken by seven or more experts (Pilot & Hunger 1999; DeVon et al. 2007). To estimate the content validity of the SQ, the researchers clearly defined the conceptual framework of spirituality by undertaking a thorough literature review and seeking expert opinion. Once the conceptual framework was established, eight purposely chosen experts in the areas of nursing, questionnaire design, and spirituality were asked to review the draft 35-item SQ to ensure it was consistent with the conceptual framework. Each reviewer independently rated the relevance of each item on the SQ to the conceptual framework using a 4-point Likert scale (1=not relevant, 2=somewhat relevant, 3=relevant, 4=very relevant). The Content Validity Index (CVI) was used to estimate the validity of the items (Lynn 1996).

2.1.2 Face validity

Face validity indicates the questionnaire appears to be appropriate to the study purpose and content area. It is the easiest validation process to undertake but it is the weakest form of validity. It evaluates the appearance of the questionnaire in terms of feasibility, readability, consistency of style and formatting, and the clarity of the language used (Haladyna 1999; Trochim 2001;   DeVon et al. 2007). Thus, face validity is a form of usability rather than reliability. To determine the face validity of the SQ, an evaluation form was developed to help respondents assess each question in terms of:

1) the clarity of the wording,

2) the likelihood the target audience would be able to answer the questions,

3) the layout and style.

Twenty five young adults with diabetes were randomly selected from two outpatient diabetes clinics and completed the face validity form on a Likert scale of 1-4, strongly disagree= 1, disagree= 2, agree= 3, and strongly agree= 4.

2.2 Construct validity

Construct validity refers to the degree to which the items on an instrument relate to the relevant theoretical construct (Kane 2001; DeVon et al. 2007). Construct validity is a quantitative value rather than a qualitative distinction between 'valid' and 'invalid'. It refers to the degree to which the intended independent variable (construct) relates to the proxy independent variable (indicator) (Hunter & Schmidt 1990). For example, in the SQ, self-awareness and meaning in life were used as proxy indicators of spirituality. When an indicator consists of multiple items, factor analysis is used to determine construct validity.

The sampling population for factor analysis was (n =160) young adults from the general population in Melbourne. The sample was recruited using a snowball sampling technique.

2.2.1 Factor analysis

Factor Analysis is a statistical method commonly used during instrument development to cluster items into common factors, interpret each factor according to the items having a high loading on it, and summarise the items into a small number of factors (Bryman & Cramer 1999). Loading refers to the measure of association between an item and a factor (Bryman & Cramer 2005). A factor is a list of items that belong together. Related items define the part of the construct that can be grouped together. Unrelated items, those that do not belong together, do not define the construct and should be deleted (Munro 2005).

Exploratory Factor Analysis (EFA) is a particular factor analysis method used to examine the relationships among variables without determining a particular hypothetical model (Bryman & Cramer 2005). EFA helps researchers define the construct based on the theoretical framework, which indicates the direction of the measure (DeVon et al. 2007) and identifies the greatest variance in scores with the smallest number of factors (Delaney 2005; Munro 2005).

It is essential to have a sufficiently large sample to enable factor analysis to be undertaken reliably (Bryman & Cramer 2005). Although, the number of participants required undertaking factor analysis remains under debate, a minimum of five participants per variable is generally recommended (Munro 2005). However, to ensure an appropriate sample size was obtained for the current study to enable factor analysis to be undertaken two criteria were considered:

1) Kaiser-Meyer-Olkin (KMO) sampling adequacy

2) Factor loadings and the correlation between a variable and a factor (Hayes 2002).

Several types of extraction methods are used to undertake factor analysis. The two most common forms are Principal Component Analysis (PCA) and Principal Axis Factoring (PAF) (Bryman & Cramer 2005). In PCA, all the variance of a variable (total variance) is analysed, while PAF only analyses common variance (Bryman & Cramer 2005). Total variance consists of both specific and common variance. Common variance refers to the variance shared by the scores of subjects with the other variables, and specific variance describes the specific variation of a variable (Bryman & Cramer 2005). Therefore, PCA is assumed to be perfectly reliable and without error (Bryman & Cramer 2005) and used on the 32 items SQ.

According to Bryman and Cramer (2005, p 330) two main criteria used to determine how many factors should be retained:

1) The Kaiser criterion to select those factors that have an eigenvalue ≥1. However, the general criterion of an eigenvalue ≥ 1.00 could misrepresent the most appropriate number of factors (Gorsuch 1983; Heppner, Lee, Wang & Park 2006).

2) A Scree Plot to depict the descending variances that account for the factors extracted in graph form. The factors that lie before the point at which eigenvalues begin to drop can be retained.

Varimax, the most commonly used orthogonal rotation was undertaken to rotate the factors to maximise the loading on each variable and minimise the loading on other variables (Field 2005; Bryman & Cramer 2005).

2.3 Reliability

Once the validity procedures were completed, the final version of the SQ was examined to assess its reliability. Reliability refers to the ability of a questionnaire to consistently measure an attribute and how well the items fit together, conceptually (Haladyna 1999; DeVon et al. 2007). Although reliability is necessary, is not sufficient to validate an instrument, because an instrument may be reliable but not valid (Beanland et al. 1999; Pilot & Hunger 1999, DeVon et al. 2007). Cronbach & Shavelson (2004) suggested researchers should consider the following issues when determining reliability:

- Standard error of the instrument, which is the most important reliability information to report.

- Independence of sampling.

- Heterogeneity of content.

- How the instrument is used.

Two estimators of reliability are commonly used: internal consistency reliability and test-retest reliability: both were used to examine the reliability of the SQ.

2.3.1 Internal Consistency Reliability

Internal consistency examines the inter-item correlations within an instrument and indicates how well the items fit together conceptually (Nunnally & Bernstein 1994; DeVon et al. 2007). In addition, a total score of all the items is computed to estimate the consistency of the whole questionnaire. Internal consistency is measured in two ways: Split-Half reliability and Cronbach's alpha correlation coefficient (Trochim 2001). In Split-Half reliability, all items that measure the same construct are divided into two sets and the correlation between the two sets is computed. Cronbach's alpha is equivalent to the average of the all possible split-half estimates and is the most frequently used reliability statistic to establish internal consistency reliability (Trochim 2001; DeVon et al. 2007).

Cronbach's alpha was computed to examine the internal consistency of the SQ. If an instrument contains two or more subscales, Cronbach's alpha should be computed for each subscale as well as the entire scale (Nunnally & Bernstein 1994; DeVon et al. 2007). Therefore, Cronbach's alpha was computed for each subscale.

2.3.2 Test-retest Reliability

Test-retest reliability is estimated by administering the same tool to the same sample on two different occasions on the assumption there will be no substantial change in the construct under study between the two sampling time points (Trochim 2001; DeVon et al. 2007). A high correlation between the scores at the two time points indicates the instrument is stable over time (Haladyna 1999; DeVon et al. 2007). The duration of time between the two tests is critical. The shorter the interval, the higher the correlation between the two tests, the longer the interval, the lower the correlation (Trochim 2001). However, very long test intervals can affect the results because of changes in participants or their environment (Linn & Gronlund 2000; DeVon et al. 2007). Currently, there is no definite evidence about the best time interval to allow between the test and the retest. Researchers need to consider factors such as the effects of time on health status such as deterioration or improvement in health and what the results will be used for, to make an appropriate decision about the time interval between tests (Concidine, Botti & Thomas 2005).

Test-Retest reliability of the SQ was undertaken by administrating the questionnaire to 25 young adults with diabetes aged 18-28, randomly selected from a diabetes outpatient clinic of a teaching hospital in an inner city area. They completed the SQ on two different occasions; at baseline and eight weeks later. Because ordinal data were obtained from the questionnaire using a four point Likert scale rated from strongly disagree to strongly agree; and the scale was not continuous, non-parametric statistical tests were deemed to be more appropriate than Pearson Correlation Coefficient (Hilton 1996; Wittkowski 2003; Jakobsson 2004). Therefore, the analysis of responses between the test and the retest was conducted using Wilcoxon Non-parametric Statistical Test to determine whether there were any significant differences between the responses at each time point.

3. Results

3.1 Translational validity

3.1.1 Content validity

According to the CVI index, a rating of three or four indicates the content is valid and consistent with the conceptual framework (Lynn 1996). For example, if five of eight content experts rate an item as relevant (3 or 4) the CVI would be 5/8=0.62, which does not meet the 0.87 (7/8) level required, and indicates the item should be dropped (Devon et al. 2007).

Therefore, three items on the draft SQ were deemed to be invalid because they yielded CVIs of 5/8=0.62 to 6/8=0.75 and were removed from the questionnaire. Those items were:

▪ 'I feel a strong connection to all people', which reviewers considered it to be similar to another item 'I have a strong emotional connection with the people around me' (CVI = 6/8=0.75)

▪ 'I respect all living creatures' (CVI=5/8=0.62).

▪ 'I accept others when they do things I think are wrong' (CVI=6/8=0.75).

All the remaining items were valid with CVIs ranging from 0.87 (7/8) to 0.100 (8/8) and were retained.

3.1.2 Face validity

All respondents rated each parameter at three or four on a Likert scale of 1-4. Ninety five percent indicated they understood the questions and found them easy to answer, and 90% indicated the appearance and layout would be acceptable to the intended target audience.

3.2 Factor analysis

To ensure having an appropriate sample size to undertake the factor analysis the KMO sampling adequacy on the SQ was 0.9. The KMO statistic varies between 0 and 1. A value of 0 indicates that the sum of partial correlations is large in

comparison to the sum of correlations, which indicates diffusion in the pattern of correlation, and that factor analysis is inappropriate. A value close to one indicates factor analysis will yield distinct and reliable factors (Field 2005). Kaiser (1974) recommended accepting values ≥ 0.5 and described values between 0.5 and 0.7 as mediocre; 0.7 and 0.8 as good, 0.8 and 0.9 as great, and > 0.9 as superb. Therefore, using Kaiser's scale, the sampling adequacy value of 0.9 for the SQ was superb. Likewise, Steven (2002) suggested that a factor is reliable if it has 10 or more variables with loadings of 0.4 and ≥150 participants. Given that the KMO of the first analysis of the draft SQ was 0.9 and all variables had loadings ≥ 0.4, the sample size of 160 was considered to be adequate to enabled factor analysis to be undertaken.

On the first run PCA, the total variance of the   draft SQ factors was 66.14%; which means at least 50% of the variance could be explained by common factors and is considered to be reasonable (Field 2005).   The communalities of the items on the SQ were > 0.5. When Kaiser's criterion was applied to the draft SQ, six factors had eigenvalues ≥1.00 in the first run PCA. A scree plot was compiled on the first PCA and indicated there were two to five factors. That is, the two tests suggested retaining a different number of factors. According to Steven (2002) and Field (2005), the scree plot and eigenvalues are accurate to determine how many factors should be retained when the sample is ≥250 and communalities (variance of the variables) are ≥ 0.6, or when the questionnaire has more than 30 variables and communalities are ≥ 0.7.

Therefore, among two to six factor solutions examined, a four factor solution with Varimax rotation was deemed to be the most statistically and conceptually appropriate to the SQ. To undertake the most appropriate interpretation, the loading values were carefully examined using Hair, Anderson, Tatham & Black's (1998) guideline for practical significance, which indicates a factor loading of ±0.3 means the item is of minimal significance, ±0.4 indicates it is more important, and ±0.5 indicates the factor is significant.

On the basis of these tests, items were eliminated from the factor pattern matrix of the SQ when the factor loading was ≤ ±0.5. The decision to eliminate such items was confirmed using Steven's (2002) Guideline of Statistical Significance for Interpreting Factor Loadings. Steven's Guideline is based on sample size and suggests that the statistically acceptable loading for 50 participants is 0.72, for 100 participants 0.51, and for 200-300 participants 0.29-0.38. The sample size used in the SQ validation process was 160: as a result, three items with a loading <0.5 were deleted. The remaining items with a loading ≥ 0.5 were accepted. One remaining item had a loading of 0.47, but was accepted because it was important to the relevant factor. The final PCA of the four-factor solution with 29 items accounted for 62.17% of the total variance. The factor loadings of the final PCA and their factorial weights are shown in Table 1.

3.3 Internal Consistency Reliability

Cronbach's alpha was computed for the revised SQ after construct validation was computed and was 0.94, which indicates a high correlation between the items and the questionnaire is consistently reliable. Opinions differ about the ideal alpha value. Some experts recommend the alpha should be at least 0.90 for instruments used in clinical settings (Nunnally & Bernstein 1994). Others suggest an alpha of 0.70 is acceptable for a new instrument (DeVellis 1991; DeVon et al. 2007). The alpha computed for each of the four subscales also exceeded the minimum value for a new tool: all subscales were ≥ 0.70, see Table 1.

3.4 Test-retest

Twenty young adults with diabetes completed the SQ in test and retest in eight weeks and Wilcoxon Non-parametric Statistical Test showed no significant differences between the two tests, see Table 2.

4. The final Spirituality Questionnaire

The final Spirituality Questionnaire includes five subscales:

1) Subscale one: "Self-awareness", which accounted for 37.11% of the total variance. This factor includes ten items and reflects information about how people view themselves. The highest loading items were: "I am satisfied with who I am" (factor loading of 0.84), "I have a number of good qualities" (loading of 0.83) and "I have a positive attitude towards myself" (loading of 0.81).

2) Subscale 2: "The importance of spiritual beliefs in life" accounted for 13.03% of variance and includes four items with very high factor loadings ranging from 0.79 to 0.82. These items refer to people's opinions about the importance of spiritual beliefs to their life.

3) Subscale 3: "Spiritual practices" accounted for 6.316% of the variance and includes six items. It focuses on people's spiritual experiences. The item "I become involved in programs to care for the environment" had the highest loading, 0.761, followed by "reading spiritual books" with a loading of 0.673, and "meditation" (0.65).

4) Subscale 4: "spiritual needs" accounted for 5.71% of the variance and includes nine items. Four items explore the search for purpose and meaning in life: "I try to find answers to the mysteries of life", "I am searching for a purpose in life" "my life is a process of becoming", "I am developing a meaningful philosophy of life", with factor loadings of 0.50 to 0.74. One item in factor 4 specifically refers to inner peace and had a loading of 0.572.

5) Two open-ended questions ask about spirituality definition and the impact of spirituality on health and wellbeing.

The items were rated on a likert scale of 1-4 where one represents strongly disagree=1, disagree=2, agree=3, and strongly agree=4.

4. Discussion

The integrity of any research depends on the accuracy of the measures used, especially when exploring complex phenomena such as spirituality. The results of the validity testing on the SQ indicated it is an accurate measure of spirituality. The processes used to validate the SQ were rigorous and appropriate. While face validity is the lowest form of validity, it was useful in that provided important information about the operationalisation of the questionnaire by young adults with diabetes. Content validity helped assess whether the content was relevant to the concept of spirituality defined for the study. Factor analysis assessed the theoretical construct of the SQ. The internal reliability (alpha) reached the recommended level for clinical use; and test-retest indicated stability of the responses to the items on the SQ over time. Therefore, the SQ could be used in routine diabetes education and management, for example, clinicians could use it confidently in usual clinical practice to incorporate spirituality into the care of their clients.

While spirituality has been recently recognised as an important aspect of health care, health care providers find it difficult to measure when they assess and care of their patients; largely because spirituality is highly subjective and often confused with religion. This paper reported the psychometric validation of the SQ to measure spirituality according to a specific definition and context: finding meaning in life, self-actualisation, and connections with the inner-self, other people and the universal whole; which are applicable to both religious and non-religious people.

However, to strengthen the rigor of the questionnaire for further research, the researchers recommend undertaking convergent and discriminant validity to examine the similarity and differences of the SQ with other spirituality tools. It is also recommended that structured equation modelling (SEM) and confirmatory factor analysis be undertaken in a larger sample with diverse healthy people as well as people with chronic illnesses to support the generalisability of the questionnaire.

5. Conclusion

The SQ is a valid and reliable research tool which can be generalised to a wider population of young people with or without diabetes.

References

Astin, A.W., Astin, H.S., Lindholm, J.A., Bryant, A.N., Szelenyi, K., Calderone, S. (2004). *The Spiritual Life of College Student, A National Study of College Students' Search for Meaning and Purpose.* Los Angeles: Higher Education Research Institute at University of California, Los Angeles.

Beanland C., Schneider Z., LoBiondo-Wood G. and Haber, J. (1999). *Nursing Research: methods, critical appraisal and utilisation.* Sydney: Mosby.

Bryman, A., Cramer, D. (1999). *Quantitative Data Analysis with SPSS Release 8 for Windows. A Guide for Social Scientists.* London: Routledge.

Bryman, A., Cramer, D. (2005). *Quantitative Data Analysis with SPSS12 and 13. A Guide for Social Scientists.* East Sussex Routledge.

Concidine, J., Botti, M., Thomas, S. (2005). Design, format, validity and reliability of multiple choice questions for use in nursing research and education. *Collegian, 12* (1), 19-24.

Coyle, J. (2002). Spirituality and health: toward a framework for exploring the relationship between spirituality and health. *Journal of Advanced Nursing, 37* (6), 589-597.

Cronbach, L.J., Shavelson, R.J. (2004). My Current Thoughts on Coefficient Alpha and Successor Procedures *Educational and Psychological Measurement, 64* (3), 391-418.

Delaney, C. (2005). The Spirituality Scale, Development and Psychometric Testing of a Holistic Instrument o Assess the Human Spiritual Dimension. *Journal of Holistic Nursing, 23* (2), 145-167.

DeVellis, R.F. (1991). Scale Development: theory and applications (Applied Social Research Methods Series, Vol. 26). Newbury Park: Sage.

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J. et al. (2007). A psychometric Toolbox for testing Validity and Reliability. *Journal of Nursing scholarship, 39* (2), 155-164.

Field, A. P. (2005). *Discovering Statistics using SPSS.* London: Sage.

Gorsuch, R. L. (1983). *Factor Analysis.* Hillsdale: NJ: Erlbaum.

Hair, J.F.; Anderson, R.E.; Tatham, R. L.; Black, W.C. (1998). *Multivariate Data Analysis*, Fifth Edition, Prentice-Hall: Upper Saddle River.

Haladyna, T. (1999). *Developing and Validating multiple-choice test items*. New Jersey: Lawrence Erlbaum.

Hayes, S.C. (2002). Acceptance, Mindfulness, and Science. *American Psychological Association, D12*, 101-105.

Heppner, P.P., Heppner, M. J., Lee, D., Wang, Y., Park, H., & W. L. (2006). Development and validation of a Collectivist Coping Style Inventory. *Journal of Counselling Psychology, 53*(1), 107-125.

Hunter, J.E., Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting errors and bias in research findings*. Newsbury Park: Sage Publications.

Kaiser, H. (1974). An index of factorial simplicity *Psychometrika, 39* (1), 31-36.

Kane, M. (2001). Current concerns in validity theory. *Journal of Educational Measurement, 38*, 319-342.

Linn, R., Gronlund, N. (2000). *Measurement and assessment in teaching*. New Jersey: Prentice Hall.

Lynn, M.R. (1996). Determination and quantification of content validity. *Nursing Research, 35*, 382-385.

Mascaro, N., Rosen, D.H., Morey, L.C. (2004). The development, construct validity, and clinical utility of the Spiritual Meaning Scale. *Personality and Individual Differences, 37*, 845-860.

McSherry, W., Draper, P. (1998). The spiritual dimension: why the absence within nursing curricula. *Nurses Education Today, 16*, 38-43.

Munro, B.H. (2005). *Statistical methods for health care research*. Philadelphia: Lippincott, Williams & Wilkins.

Nunnally, J.C., Bernstein, I.H. (1994). *Psychometric theory*. New York: McGraw-Hill.

Pilot, D., & Hunger, B. (1999). *Nursing research: principals and methods*. Philadelphia: Lippincott Williams & Wilkins.

Stevens, J. (2002). *Applied Multivariate Statistics for the Social Sciences (4th Edition)*. Mahwah, NJ: Lawrence Erlbaum Associates.

Tanyi, R. A. (2002). Towards clarification of the meaning of spirituality. *Advanced Nursing, 39* (5), 500-509.

Trochim, W.M.K. (2001). *The Research Methods Knowledge Base*. Cincinnati: Atomic Dog.

Tse, S. et al. (2005). Exploration of Australian and New Zealand indigenous people's spirituality and mental health. *Australian Occupational Therapy Journal, 52* (3), 181-187.

Tuck, I., Wallace, D., & Pullen, L. (2001). Spirituality and Spiritual Care Provided by Parish Nurses. *Western Journal of Nursing Research, 23* (5), 441-453.

Underwood, L.G., Institute, F., Teresi, J.A. (2002). The daily Spiritual Experience Scale: Development, theoretical description, Reliability, Exploratory factor analysis, and Preliminary Construct Validity Using Health-Related Data. *Annals of behavioural medicine, 24* (1), 22-33.

Table 1. The results of the final four factor solution of the SQ according to the Principal Component Analysis with Varimax rotation and the internal consistency of each factor

Items	Factor 1	Factor 2	Factor 3	Factor 4
Self-Awareness (α=0.91)	*Loadings*			
I am satisfied with who I am.	0.84			
I have good qualities.	0.83			
I have positive attitude towards myself.	0.81			
I am a worthy person.	0.80			
I am generally self-confident.	0.69			
My life in meaningful to me.	0.64			
I believe I am equal to other people.	0.62			
I am a compassionate person.	0.61			
I find meaning in difficult situations.	0.58			
I think about positive things.	0.55			
Importance of spiritual beliefs in life (α=0.91)		*Loadings*		
My spirituality helps me define my goals.		0.82		
My spirituality helps me decide who I am.		0.82		
My spirituality is part of my whole approach to life.		0.81		
My spirituality is integrated into my life.		0.80		
Spiritual practices (α=0.80)			*Loadings*	
I become involved in environmental programs.			0.77	
I read books about spirituality.			0.67	
I meditate to achieve inner peace.			0.65	
I try to live in harmony with nature.			0.57	
I try to find any opportunity to enhance spirituality.			0.55	
I use silence to get in touch with myself.			0.47	
Spiritual needs (α=0.89)				*Loadings*
I search for a purpose in life.				0.74
I enjoy listening to music.				0.69
I need to find answers to the life mysteries.				0.66
Maintaining relationships is important to me.				0.64
I need to attain inner peace.				0.57
I seek beauty in my life.				0.55
I need to have strong connections with people.				0.55
My life in evolving.				0.55
I need to develop a meaningful life.				0.50

Table 2. Test-Retest results of the SQ using Wilcoxon's non-parametric test. As the table indicates, there were no significant differences in the P values at the level of 0.05 in the responses to the items between the two tests.

| Items	Asymp. Sig. (P value)
I am satisfied with who I am.	0.157
I have good qualities.	0.480
I have positive attitude towards myself.	0.083
I am a worthy person.	0.705
I am generally self-confident.	0.655
My life in meaningful to me.	0.180
I believe I am equal to other people.	0.705
I am a compassionate person.	0.180
I find meaning in difficult situations.	0.157
I think about positive things.	1.000
My spirituality helps me define my goals.	0.414
My spirituality helps me decide who I am.	1.000
My spirituality is part of my whole approach to life.	0.317
My spirituality is integrated into my life.	0.180
I become involved in environmental programs.	0.234
I read books about spirituality.	0.655
I meditate to achieve inner peace.	0.655
I try to live in harmony with nature.	0.317
I try to find any opportunity to enhance spirituality.	0.096
I use silence to get in touch with myself.	0.157
I search for a purpose in life.	0.763
I enjoy listening to music.	0.157
I need to find answers to the life mysteries.	0.564
Maintaining relationships is important to me.	0.564
I need to attain inner peace.	0.564
I seek beauty in my life.	0.157
I need to have strong connections with people.	1.000
My life in evolving.	0.783
I need to develop a meaningful life.	0.180
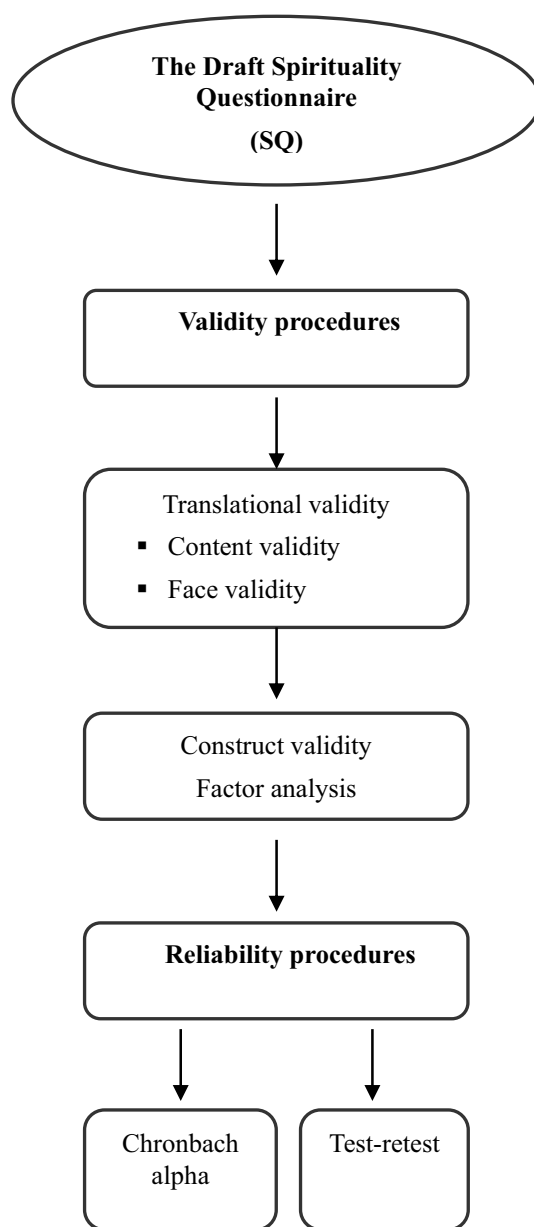
Figure 1. A flow chart depicting the process used to validate the Spirituality Questionnaire (SQ).