Diagnostic validity of the anxiety and depression questions from the Well-Being Process Questionnaire

Gary Williams[1] and Andrew P Smith[1]
Centre for Occupational & Health Psychology, School of Psychology, Cardiff University, 63 Park Place, Cardiff, CF10 3AS

Corresponding author:
Andrew P Smith
Tel: +44(0)29 208 76574
E-mail address: smithap@cardiff.ac.uk

Abstract

Background: Previous research shows that the Well-being Process Questionnaire (WPQ) has good content validity, construct validity, discriminant validity and reliability.

Objective: The present research examined the diagnostic validity of the anxiety and depression questions from the WPQ by comparing them with the Hospital Anxiety and Depression scale (HADS) from which they were derived.

Method: One hundred and twenty university staff members aged 20-64 participated in the study which involved an anonymous online survey. The data were used to assess the ability of single item measures, rated on a 10 point scale, to correctly identify a respondent that meets a diagnostic criteria, in this case clinical levels of depression or anxiety.

Results: This analysis involved comparison with an established measure (HADS clinical cut-off) in terms of the proportion of those with the condition correctly identified as such (sensitivity) by the single items and the proportion without the condition correctly identified as such (specificity) by the single items. The results showed that a cut-off point at a score of 5 provided the best results for sensitivity and specificity in the depression and anxiety items. Sensitivity at this point was 71.4% and 86.3% for depression and anxiety respectively, while specificity was 85.4% for depression and 72.6% for anxiety.

Conclusion: These findings confirm that the single item anxiety and depression questions from the WPQ can be used as an initial screening tool to identify clinical cases of anxiety and depression.

Relevance for patients: This will provide a rapid method of assessment that will benefit patients and lead to more effective prevention and management.

Keywords

1. Introduction

The Well-Being Process Questionnaire (WPQ) is a measuring instrument that uses short versions of established questionnaires to investigate factors which change well-being and well-being outcomes. Choice of variables has been linked to the Demands-Resources-Individual effects model [1] which has been validated using the original scales in studies with nurses and teachers [2-3]. Similar results have been obtained using the WPQ [4-6]. The aim of the present article is to focus on the anxiety and depression questions from the WPQ and examine their diagnostic validity. Diagnostic validity is one aspect of the psychometrics of the WPQ and it is important to initially briefly review other aspects of validity and reliability of the WPQ (see [7] for a more detailed account)

The validity of a measure is assessed in relation to its intended purpose [8,9]. Content and construct validity, which are concerned with how well the measure represents the variable for which it is designed, are of primary importance for outcome measures because the purpose of these measures is to accurately represent the variable in question [9]. Predictive validity is more important for predictor variables [8,9], as their purpose is to predict scores on a particular outcome. The content validity of measures should be ensured in the construction of the measure itself [9]. In previous research this has been achieved by sampling items from a pool of already available items [10]. In the present research a similar approach was followed by sampling questions from the multi-item measures into the examples provided alongside the single-item measures of the WPQ. Measures of abstract variables require construct validity, which is concerned with how well the results represent the behaviour of a person with the construct in question [8], therefore providing a proxy measure of a variable that cannot be measured directly. One approach to this type of validation in a newly-designed measure is to compare it to an established measure of the same variable, referred to as concurrent validity when the measures are completed at the same time [8]. The single-item measures from the WPQ were compared to established multi-item measures already in use, in order to establish their construct or concurrent validity. From a statistical standpoint, correlations of .80 are used to consider two sets of scores as unidimensional [11]. However, correlations between two measures designed to assess the same construct are frequently lower due to attenuation [9]. For example, Costa and McCrae [12] found correlations of .65 for two extraversion measures and .68 for two neuroticism measures (Sixteen personality factor questionnaire and Eysenck personality inventory). This research suggests that moderate correlations are often found between accepted multi-item measures of the same construct and these standards can also be applied to single-item measures. The correlations between the anxiety and depression single items from the WPQ and the anxiety and depression scales from the HADS were both 0.65, suggesting good validity [13]. In addition, the single anxiety and depression items were shown to be sensitive outcomes of established predictor variables (work characteristics and personality, [4].

Reliability is possibly the most commonly criticized, yet infrequently assessed, aspect of single-item measures [14]. This is most likely due to an assumed inability to estimate the internal consistency of a single-item, although multiple approaches do exist. One approach is based on the correlations between scores taken at 3 points in time [15] (Equation 9), however the most direct approach is using the Wanous [16,17] method, using the correction for attenuation formula. Using this formula, Wanous et al [17] estimated the minimum reliability of a single-item job satisfaction question as .70 based on a meta-analysis of 17 studies. The estimated reliability (using the Wanous method) of the single item anxiety score was 0.66, and the reliability of the single item depression score was 0.61.

The aim of this article is to examine the diagnostic validity of the single item anxiety and depression variables in the WPQ. Diagnostic validity refers to the ability of the measure to correctly identify a respondent that meets a diagnostic criterion, for example clinical levels of depression or anxiety. This method involves comparison with an established measure in terms of the proportion of those with the condition correctly identified as such (sensitivity) and the proportion without the condition correctly identified as such (specificity). This is most common in depression and anxiety research where specific

criteria exist for diagnosis of clinical depression or anxiety. This is relevant to our aim of developing short, practical measures and the newly designed measures should identify those with high or low mental health accurately so that, when the measures are applied in practice, intervention is targeted at the right people. In the present study, the single-item depression and anxiety measures were compared against the HADS measures, which contain cut-off scores for cases of depression or anxiety and a 'normal range'.

2. Method

This research was approved by the Ethics committee, School of Psychology, Cardiff University, and carried out with the informed consent of the participants. Participants were recruited from the staff of Cardiff University using an advertisement on the university noticeboard. The study involved an online survey presented using SurveyTracker that they could complete in their own time. Participants were instructed that they could skip any questions that they were not comfortable answering, and all data were provided anonymously. Informed consent was achieved within the questionnaire where participants without agreeing could not continue beyond the consent page. Following the consent page, participants were presented with an instructions sheet and a debrief sheet.

2.1. Participants

One hundred and twenty university staff members aged 20-64 participated in the study. Participants from all areas of the university were able to participate, including finance, teaching, accommodation, and security, although the role of specific respondents was not recorded. The majority were aged 30-39 (32%), married or living with a partner (63%) and were educated to degree or higher degree level (73%). Working patterns were most commonly full-time (81%) fixed hours (79%). This sample was considered representative of a typical UK university and the number in the sample allowed identification of a medium effect size

2.2. Materials

The survey consisted of single items derived from established psychosocial scales. The complete questionnaire is described in detail in [6]. The key questions in the present analyses are the anxiety and depression variables which are shown below:

Anxiety: On a scale of one to ten, how anxious would you say you are in general? (e.g. feeling tense or 'wound up', unable to relax, feelings of worry or panic)?

1= Not at all 10 = Extremely

Depression: On a scale of one to ten, how depressed would you say you are in general? (e.g. feeling 'down', no longer looking forward to things or enjoying things that you used to)?

1= Not at all 10 = Extremely

The scores on these variables were compared with the anxiety and depression scores from the Hospital Anxiety and Depression Scale [18].

2.3. Analysis strategy

The diagnostic validity of the outcome measures was examined by creating groups based on the distribution of scores on outcome measures using the single- and multi-item scales. Sensitivity and specificity were examined for depression and anxiety where the multi-item scale includes known cut-offs for cases of clinical depression or anxiety. Low and high scores based on the possible range on the single items for depression and anxiety were compared with established cut-offs for 'no depression/anxiety' and 'mild' to 'severe' depression/anxiety.

3. Results

Table 1 shows the demographic charateristics of the sample.

Table 1. Demographic characteristics of the sample

Age	Mean: 36.9 years s.d. 10.6 range 21-64 years
Gender	Male: 27.5%
Marital status	Single: 31.7% Living with partner: 25.8% Married: 36.7% Separated/Divorced: 5.8%
Education:	To degree level: 73.3%
Ethnicity	White/British: 98.3%

Correlations between the short and long versions of the questions were similar for the different demographic groups (e.g. married: depression r = 0.61, anxiety r = 0.49; single: depression r= 0.66, anxiety r = 0.49 . Educated to degree level: depression r = 0.66, anxiety r = 0.51. Not educated to degree level: depression r = 0.51, anxiety r = 0.54).

Table 2 shows the sensitivity and specificity for the WPQ single item depression and anxiety values plotted against normal and moderate/severe categories from the HADS. A cut-off point at a score of 5 provided the best results for sensitivity and specificity in the depression and anxiety items. Sensitivity at this point was 71.4% and 86.3% for depression and anxiety respectively, while specificity was 85.4% for depression and 72.6% for anxiety.

Table 2. Sensitivity and specificity for single-item measures compared to multi-item measures of depression and anxiety for each cut-off score on the single-item measure (bold=median score).

Depression			Anxiety		
Cut-off point	Normal range (1-7)	Mild-Severe (8+)	Cut-off point	Normal range (1-7)	Mild-Severe (8+)
9	-	-	9	98.4	0
8	100	4.8	8	98.4	11.8
7	95.1	19	7	98.4	35.3
6	91.5	33.3	**6**	**85.5**	**64.7**
5	85.4	71.4	5	72.6	86.3
4	78	76.2	4	62.9	90.2
3	**65.9**	**90.5**	3	48.4	98
2	43.9	90.5	2	32.3	100
1	8.5	100	1	8.1	100

4. Discussion

The present results show that single items measuring anxiety and depression have good diagnostic validity. This, combined with good content validity, predictive validity and reliability, suggests that the WPQ can be

used in applied situations such as an initial screening tool for mental health problems. This initial screening could even be done remotely (e.g. over the telephone or by e-mail) to reduce logistic problems associated with primary care appointments. Similarly, it has other practical features, such as being low cost, that make it a useful alternative to established methods. However, it should be emphasised that this initial screen would then lead to a more detailed consideration of the causes and prevention/management of the mental health problem. The WPQ, based on the adapted DRIVE model, can now be used in longitudinal studies and to evaluate interventions. These interventions could involve changes in job characteristics, development of coping skills or therapeutic approaches dealing with established problems. The presence of an underlying model and short measuring instrument will enable more effective prevention and management of negative influences and outcomes, and also allow promotion of positive well-being.

The present study has it's limitations, most of which reflect the sample used. Future studies should investigate a sample which is more representative of the general population and also a sample of clinical patients. It is important to know whether similar results are obtained with the single items with groups who have less awareness of the concepts of anxiety and depression.

5. Summary

In summary, the present study has shown that the WPQ single items measuring anxiety and depression have good diagnostic validity. The use of these items enables one to assess common mental health problems very quickly and the WPQ is an ideal tool to use in both audits of psychosocial factors and the assessment of interventions.

References

1.   Mark, G.M. & Smith, A.P. 2008. Stress models: A review and suggested new direction. In: Occupational Health Psychology: European Perspectives on research, education and practice. Vol. 3. EA-OHP series. Edited by J.Houdmont & S. Leka. Nottingham University Press. 111-144.

2.   Mark, G & Smith, A.P. 2011. Effects of occupational stress, job characteristics, coping and attributional style on the mental health and job satisfaction of university employees. Anxiety, Stress and Coping, 25, 63-78. Doi: 10.1080/10615806.2010.548088

3.   Mark, G. & Smith, A.P. 2012. Occupational stress, job characteristics, coping and mental health of nurses. British Journal of Health Psychology, 17, 505-521. Doi: 10.1111/j.2044-8287.2011.02051.x

4.   Williams, G.M. & Smith, A.P. 2016. Using single-item measures to examine the relationships between work, personality, and well-being in the workplace. Psychology: Special Edition on Positive Psychology, 7, 753-767.

5.   Williams, G., Pendlebury, H. & Smith, A.P. 2017. Stress and Well-being of Nurses: an Investigation using the Demands-Resources- Individual Effects (DRIVE) model and Well-being Process Questionnaire (WPQ). Jacobs Journal of Depression and Anxiet,1,1-8..

6.   Williams, G., Thomas, K & Smith, A.P. 2017. Stress and Well-being of University Staff: an Investigation using the Demands-Resources- Individual Effects (DRIVE) model and Well-being Process Questionnaire (WPQ). Psychology, 8, 1919-1940. https://doi.org/10.4236/psych.2017.812124

7.   Williams, G.M.. 2015. Researching and developing mental health and wellbeing assessment tools for supporting employees and employers in Wales. PhD thesis, Cardiff University. http://orca.cf.ac.uk/71443/1/2015williamsphd.pdf

8.   Cronbach, L. J. (1990). Essentials of Psychological Testing (5th ed.). New York, NY: HarperCollinsPublishers, Inc.

9.   Nunnally, J. C. (1978). Psychometric theory (2nd ed.). London: McGraw-Hill.

10.   Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the Positive and Negative Affect Schedule (PANAS). Journal of Cross-Cultural Psychology, 38(2), 227-242.

11.   Tabachnick, B. G., & Fidell, L. S. (2007). Using Multivariate Statistics (5th ed.). Boston, MA: Pearson Education, Inc.

12.   Costa, P. T., & McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. Journal of Personality and Social Psychology, 38(4), 668-678.

13.   Williams, G. & Smith, A.P. 2013. Measuring wellbeing in the workplace: Single item scales of depression and anxiety. In Contemporary Ergonomics and Human Factors 2013. Martin Anderson (ed). CRC Press: Taylor & Francis. London. ISBN 978-1-138-00042-1. Pg 87-94.

14.   Viswanathan, M., Bergen, M., Dutta, S., & Childers, T. (1996). Does a single response category in a scale completely capture a response? Psychology and Marketing, 13(5), 457-479.

15.   Heise, D. R. (1969). Separating reliability and stability in test-retest correlation. American sociological review, 93-101.

16.   Wanous, J. P., & Hudy, M. J. (2001). Single-Item Reliability: A Replication and Extension. Organizational Research Methods, 4(4), 361-375.

17.   Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: How good are single-item measures? Journal of Applied Psychology, 82(2), 247-251.

18.   Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. Acta Psychiatrica Scandinavica, 67(6), 361-370.