The Fear-Avoidance Beliefs Questionnaire (FABQ): A Critique of Instrument Development

Mary Margaret Williamson

Dr. Jessica Wallace

The University of Alabama

Department of Health Sciences

Southeastern Universities Graduate Research Symposium 2023

Abstract

The purpose of this paper was to evaluate the development of a measure for fear-avoidance beliefs commonly used in clinical practice: the fear-avoidance beliefs questionnaire (FABQ). The FABQ was created to evaluate risk for prolonged recovery due to chronic low back pain by measuring elevated fear-avoidance beliefs related to physical activity and work. The FABQ has been used in various clinical settings to aid healthcare providers in making clinical decisions regarding patient care after injury and/or diagnosis of chronic pain. Given the gravity of clinical implications of inferences made from the use of this scale, thorough analysis of the psychometric development of the FABQ is vital for informing clinicians who seek to use this fear-avoidance measure. There were strengths and limitations present in the questionnaire's developmental methodology and are discussed in the present critical review.

The Fear Avoidance Beliefs Questionnaire (FABQ): A Critique of Instrument Development

Avoidance learning was a theory proposed by Fordyce et al. (1982) as a process of operant conditioning where anticipated consequences to painful stimuli encourage avoidant beliefs to escape, thus reinforcing the avoidant behavior. Further evidence indicated that there was little support that these avoidance behaviors provide a reduction of chronic pain (Philips 1987). A theoretical model of fear-avoidance was developed by Lethem at al. (1983) to describe the cyclical relationship that these described fear avoidance beliefs (FAB) and behaviors have with other biopsychosocial factors and how those relationships interact to perpetuate chronic pain. A psychometric measure was needed to assess risk for elevated FAB, particularly in patients with chronic low back pain (LBP) that impacted their ability to work and be physically active. Other psychometric measures existed at the time; however, they were not designed to assess attitudes and beliefs related to work ability.

Scale Description

Therefore, the authors main objective was to develop a scale that assesses an individual's beliefs related to the effects their work and physical activity have on their chronic LBP. They developed the Fear Avoidance Beliefs Questionnaire (FABQ) by applying fear theory, ideas about fear-avoidance cognitions, somatic awareness and focusing, and concepts of disease conviction like illness severity beliefs and how the illness affects an individual's life.

Scale Development

Item Selection

First, authors conducted a literature review to identify a need for a new scale. Similar scales existed at the time, but none that purposely measured FAB about pain related to work. The authors then developed 10 original items based on Fordyce's observations about individual's pain behaviors (Fordyce et al. 1968, Fordyce 1976, Fordyce et al. 1982); however, they do not go into depth on what observations specifically lead to which items. In fact, there is very little explanation as to how the original items were developed and no formal panel of experts was used to establish content validity. The final six items came other existing scales. Two items were derived from a study by Sandsrtom and Esbjornsson (1986) where participants were interviewed about their return to work following a rehabilitation program for LBP. Two statements were turned into items, though it is not explicitly said which items were derived from these statements. Three items were similar to items in the harm subscale of the Survey of Pain Attitudes (SOPA), and one was similar to items in the Pain and Impairment Relationship Scale (PAIRS) Beliefs Questionnaire. Again, the authors neglected to provide which items were used to develop the remaining four items. All 16 items of the FABQ were scored on a 7-point Likert scale ranging from zero (strongly disagree) to six (strongly agree). There is disagreement between what the authors stated in the article and the copy of the FABQ items listed in the Appendix, as the Appendix lists items ranging from zero (completely disagree) to six (completely agree). This change in diction could have significant implications to their results' internal and external validity if different versions were used throughout the pilot and main studies' testing or if their studies used "strongly" but other studies or clinicians utilize the "completely" version listed in the Appendix.

Samples

Authors indicate that there are five pilot studies involved in the creation of the FABQ, but there is little direction within the article to indicate distinct studies. It is also possible that the four other pilot studies were included in the introduction, but again they were not clearly identified. In the presented study, four samples can be identified from the text: a sample of n=30 orthopedic out-patient clinic patients segregated into two pilot groups used to develop the FABQ, n=26 hospital physiotherapy out-patients with LBP used to assess reliability of the FABQ, n=184 LBP patients with and without sciatica used for the main study assessing validity of the FABQ and its relationship to other outcomes related to chronic LBP, and finally a sample of n=210 patients combining the 26 physiotherapy and 184 hospital out-patient patients was randomly divided into two groups and used to assess factor structure. The extent of overlap between samples is not clearly defined, nor how the samples were acquired. No mention of a power analysis was made, and despite the sample for factor structure being somewhat adequate (Nunnally & Bernstein 1994, Pett, Lackey, & Sullivan 2003), a larger sample size would have likely resulted in a more stable scale and a better understanding of the factor structure (Tabachnick & Fidell 2001, Comrey, 1973). The authors reported that all participants were English-speaking Caucasians with a mean age of 39.7 ±11.7 (male 55.7%, female 44.3%, age range 18-60 years old). The mean time between initial bout of LBP and initial assessment for the study was 7.4 ±8.4 years, and the mean duration of the most current bout of LBP for participants was 13.7 ±19.8 months. Only 34% of participants were still working, with 24% being employed but off from work, 15% who were unemployed, and 27% where work status was irrelevant. For those off work, the

mean time between when they last worked and study initial assessment was 4.0 ±4.3 months, and total time off of work within the previous year was 6.0 ±10.8 months. The time period between initial assessment and the first therapy session for the 26 out-patient participants was 48 hours and was chosen so that no treatment was provided between sessions that would interfere with clinical or cognitive outcomes while also decreasing the likelihood that participants would be able to recall items. Patients with spinal pathologies, histories of psychiatric disorders or alcoholism, or who were illiterate were excluded from all samples, so it is unclear how many participants may have been included in initial sampling.

Data Analysis

To assess test-retest reliability of the FABQ, simple percent agreement was assessed using *K* statistics to provide allowances for distribution effects and level of chance agreement. To assess factor structure, a principal-components analysis (PCA) with varimax orthogonal rotation was used in the exploratory analysis. Since the FABQ is based on the Fear Avoidance Theory of Chronic Pain, perhaps an oblique rotation should have been run first before moving on to an orthogonal rotation. Also, a common factor analysis would have been more appropriate to run prior to a PCA due to the novelty of the scale. Once true factor structure was established, then a PCA could be run (Costello & Osborne 2005). Final factor model structures were chosen based on scree, mineigen, item analysis and correlation matrix, with items being accepted in the final factor(s) if they loaded at least 0.45 on a factor, but less than 0.30 on another factor. A non-orthogonal rotation was used to assess for a higher order factor(s), but the rotation was not explicitly defined. Regression analysis was used to investigate

relationships between the FAB and pain severity, total pain duration, current pain episode duration, disability related to ADL, disability related to current work loss, disability related to work loss within the last year, and scores on the Modified Somatic Perception Questionnaire (MSPQ) and Zung Depressive Inventory (ZDI).

Validity

Significant relationships between FAB and disability in the LBP sample were evaluated via a hierarchical multiple linear regression analysis. The authors decided on a fixed order of entry in the regression analysis based on the theory that LBP is the primary problem and should be controlled for before including FAB related to work and physical activity, and finally depressive symptoms which are thought to develop secondarily to pain.

Results

Item Distribution

No items required removal due to lack of participant understanding or compliance on the scale; however, 21 participants were not employed and thus could not complete items 6-16. None of the items had greater than 2.1% missing answers and only 1.3% of answers were missing overall. No items were excluded due to excessive skewness, as all answers on all items fell between at least four categories. Values for skewness and data normality was not provided for the data, and lack of excessive skewness was the only mention of normality made.

Reliability

All 16 items were found to be acceptably reliable, with 71% of participants answering identically on the test and retest. All items were found to have a high level of reproducibility according to the *K* statistics (mean K = 0.74, sig. < 0.001). Six items were found to have nearly complete concordance (>0.80), eight items had substantial concordance (0.61-0.80), and only two items had moderate concordance (0.41-0.60). It is unclear whether or not reliability was assessed prior to validity, but its inclusion prior to discussion of factor and validity analysis indicates the possibility. The authors indicated that the Pearson product-moment correlation coefficients for two scales were 0.95 and 0.88. It is not clear what two scales they meant, though it can be inferred that the correlation coefficients are that of the test and retests. Clarity here, especially when Pearson product-moment correlation coefficients are used to establish criterion validity, is needed as these statistics were included with reliability results.

Factor Analysis

A principal-components analysis identified a 2-factor structure as the best fit for the scale over a single or 3-factor structure. The 3-factor structure was dismissed due to instability from the third factor barely reaching mineigen criteria and only accounting for a “small percentage of the variance”. No statistical output was provided for the 3-factor structure, nor was the percentage of variance defined, which would have been valuable to view since the factor evidently met criteria. Alternatively, the single-factor structure, which excluded items 1, 8, 13, 14, and 16, had an α value of 0.82. Despite this, the 2-factor structure was favored due to scree, mineigen, correlation matrix, and the intention for the scale to have two dimensions in FAB related to work and physical activity. The two factors were identified, excluding the same items as the single-factor structure, and

was supported by non-orthogonal analysis. Item analysis found item 1 to have low communality and inconsistent factor loading, a failure of item 8 to load on either factor, and that items 13, 14 and 16 were redundant. Factor 1, made up of items 6, 7, 10, 11, 12 and 15, focused on FAB related to the relationship between work and LBP, while Factor 2 focused on FAB related to physical activity generally (items 2, 3, 4, 5). Both factors intercorrelate (0.39). Despite their failure to be included in either factor or in the single-factor structure, items 1, 8, 13, 14 and 16 were still included in the final scale. An explanation was not provided by the authors as to why they were included. Their motivations could have been useful in understanding the clinical utility of leaving these un-factored items, especially since items 13, 14 and 16 were found to be redundant.

Validity

The work factor was weakly related to sex (male 24 $\pm$12.8, female 19.3 $\pm$12.7, *t* value 2.39, $P = 0.02$) while the physical activity factor was not, and neither correlated with age. While not correlating strongly with most biomedical measures of pain, the work factor correlated strongly with disability related to activities of daily living (ADL, 0.55, $p < 0.001$), disability related to current loss of work (0.39, $p < 0.001$), disability related to lost work from the past year (0.55, $p < 0.001$), MSPQ scores (0.36, $p < 0.001$), and ZDI scores (0.41, $p < 0.001$). This indicates a strong relationship between FAB related to work and psychosocial constructs like disability and depression. The physical activity factor had similar, though weaker, relationships with measures of disability (ADL 0.51, $p < 0.001$; work loss in past year 0.23, $p < 0.01$) and depression (depressive symptoms 0.36, $p < 0.001$). The work factor explained 23% of the variance related to disability in ADL and 26% of the variance in work loss in the past year. The

physical activity factor only explained an additional 9% variance for ADL. Males in the sample demonstrated a stronger relationship between work-related FAB and disability (σ = 0.27) and lost work in the previous year (σ = 0.35), while females accounted for 17% and 18% of the variance. The authors indicated that a gender and beliefs interaction was not significant after entering gender into the regression equation. Having been written in the 1990s, the mixing of the terms “sex” and “gender” was more common, but an analysis of both in the future is needed to determine their individual relationships with FAB in LBP patients.

Strengths

The FABQ is potentially a clinically useful instrument for measuring FAB in workers suffering from chronic LBP due to its brief nature, strong test-retest reliability, and the significant positive relationship between the work subscale and disability and work status change.

Limitations

While the SOPA is a long measure, an assessment of the concurrent or convergent validity between it and the FABQ would have been wise considering the closeness of their respective natures. The authors also indicated that their results needed to be checked and compared to other validated measures, thereby establishing construct and criterion validity. The study did not examine reliability or predictive validity in those experiencing acute LBP and was only found to have adequate reliability and validity in a sample of employed or recently employed individuals with chronic LBP. The authors asserted that their findings established the FABQ as reliable and valid in this population, which is generous considering their sample size. They also did not

determine how long an individual had to be unemployed before there was reduced utility in the application of the work subscale.

The authors did not go into sufficient detail about the item development process, and this is particularly concerning as it calls into question the content validity of the scale and its subscales. If a panel of experts was used, that process should have been described, as well as a formal power analysis described for selecting an adequate sample size.

Most importantly, this study suffered from a lack of clarity in many areas where elaboration would have lent transparency, intelligibility and thoroughness, including in the explanation of the pilot studies, how the items were developed, the factor structure-related statistics (scree, mineigen, correlation matrix, item analysis), and for why the non-factor loading items were allowed to remain.

Conclusion and Future Research

The FABQ was found to be a reliable and valid measure of FAB in a sample of chronic LBP patients, and the work subscale demonstrated a strong positive relationship with work loss and disability in ADL. The authors indicated further research is needed to compare FABQ and PAIRS and SOPA. A cursory search of the available literature indicates this has yet to be done despite the close nature of the FABQ and SOPA. Predictive and discriminant validity need to be established for the FABQ for it to be clinically relevant and easily used by clinicians. Much larger samples are needed to further assess the reliability and validities of the FABQ in both chronic LBP patients and patients with other chronic and acute pain conditions that affect work and physical activity.

References

Comrey, A. L. (1973) *A first courses in factor analysis*. New York, Academic Press.

Costello, A. B., and Osborne, J.W. (2005) Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical Assessment Research & Evaluation, 10(7): 1-9.

Fordyce, W.E., Fowler, R.S., Lehmann, J.F., and Delateur, B.J., Some implications of learning in problems of chronic pain. J. Chron. Dis., 21 (1968) 179-190.

Fordyce, W.E. (1976) *Behavioral methods for chronic pain and illness*, Mosby, St. Louis, MO.

Fordyce, W.E., Shelton, J.L., and Dundore, D.E. The modification of avoidance learning pain behaviors, J. Behav. Med., 5 (1982) 405-414.

Lethem, J., Slade, P.D., Troup, J.D.G., and Bentley, G. Outline of a fear-avoidance model of exaggerated pain perception. Behav. Res. Ther., 21 (1983) 401-408.

Nunnally, J.C., & Bernstein, I.H., (1994) *Psychometric theory (3rd ed.).* NY: McGraw-Hill.

Pert, M., Lackey, N., & Sullivan, J., (2003) *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Thousand Oaks, CA: Sage Publications.

Philips, H.C. Avoidance learning and its role in sustaining chronic pain. Behav. Res. Ther., 25 (1987) 273-279.

Sandstrom, J., and Esbjornsson, E., Return to work after rehabilitation: The significance of the patient's own prediction, Scand. J. Rehab. Med., 18 (1986) 29-33.

Tabachnick, B. G. and L. S. Fidell. (2001) *Using multivariate statistics*. Needham Heights, MA, Allyn & Bacon.