Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

European Journal of Pain 11 (2007) 341–351

www.EuropeanJournalPain.com

Cognitive assessment of musculoskeletal pain with a newly validated Greek version of the Fear-Avoidance Beliefs Questionnaire (FABQ)

George Georgoudis [a,b,*], George Papathanasiou [b], Panagiotis Spiropoulos [b], Kostantinos Katsoulakis [c]

[a] *Centre for Rehabilitation Science, University of Manchester, Manchester Royal Infirmary, Oxford Road, Manchester, UK*
[b] *Division of Musculoskeletal System, Physical Therapy Department, Technological Educational Institution of Athens, Greece*
[c] *'Agia Sofia' General Hospital, Athens, Greece*

Received 27 August 2005; received in revised form 5 April 2006; accepted 1 May 2006
Available online 13 June 2006

Abstract

Fear of pain and avoidance are psychological factors of primary importance when assessing chronic musculoskeletal pain, which are often measured with the Fear-Avoidance Beliefs Questionnaire (FABQ). Both two- and three-subscale versions have been described. The aims of this study were: to assess the cognitive traits of musculoskeletal pain patients using a newly validated Greek version of the FABQ, and to further examine the construct validity and responsiveness of the measure. Factor analysis yielded three factors that accounted for 65% of the total variance. Physical activity explained 12.3% of the variance and was identical to the original version, unlike the work subscale which split into two: the FABQ work1 related to "work as cause" (15.2% of the variance) and the FABQ work2 related to "work as prognosis" (37.5% of the variance). Internal consistency was good (0.72–0.90). Test–retest reliability was satisfactory and close to the original version both for individual items and the subscales. Responsiveness of the 3-factor model was satisfactorily assessed as the ability to detect: (A) change in general – (paired *t* test, effect size); (B) clinically important change (paired *t* test, standardised effect size), and (C) real change in the concept being measured (ROC analysis). Construct validity of the FABQ was shown through the interaction with anxiety and depression, pain control and responsibility, psychological distress and pain intensity, and criterion-related validity through the association with another fear-avoidance measure (TSK). New aspects of responsiveness and construct validity were demonstrated for the FABQ, using a three-subscale validated Greek version.

Keywords: FABQ; Greek; Questionnaire; Pain; Validity

1. Background

The importance of fear-avoidance in the development of chronic pain is undoubted and supported by the vast number of reports being published. In a systematic review to identify the risk factors of chronicity in spinal pain, Linton provided Level A evidence demonstrating fear-avoidance beliefs to be related to pain and disability (Linton, 2000).

Based on cognitive–behavioural theories (Lethem et al., 1983; Vlaeyen and Linton, 2000; Waddell et al., 1993), fear of pain and avoidance can lead to decreased self-efficacy, further avoidance, deconditioning and disability. On this conceptual background, two self-report questionnaires have been developed to quantify the fear-avoidance beliefs: the Tampa Scale of Kinesiophobia

* Corresponding author. Present address: 106 Antoniou Mpertou Str., Amfiali, Pireas, Greece. Tel.: +30 210 4014920.
E-mail address: gg@hol.gr (G. Georgoudis).

1090-3801/$32
doi:10.1016/j.ejpain.2006.05.001

(Kori et al., 1990) and the Fear-Avoidance Beliefs Questionnaire (Waddell et al., 1993).

In 1993, Waddell et al. developed a self-report questionnaire assessing fear-avoidance beliefs regarding the effects of physical activities and work on low back pain and investigated the psychometric properties of the questionnaire on a group of patients with chronic low back pain (Waddell et al., 1993). Results from the study indicated that the FABQ was a reliable measure and two subscales within the FABQ were identified: a four-item scale measuring fear-avoidance beliefs about physical activity (FABQ/physical) and a seven-item scale measuring fear-avoidance beliefs about work (FABQ/work) (Waddell et al., 1993). The two FABQ subscales have good internal consistency ranging from a 0.84 to 0.92 on the work subscale and a 0.52 to 0.77 on the physical activity subscale (Crombez et al., 1999; Waddell et al., 1993). Both subscales have good test—retest stability (Pfingsten et al., 2000; Waddell et al., 1993) and interact significantly with correlation varying from $r = 0.39$ to 0.60 (Crombez et al., 1999; Waddell et al., 1993). The TSK and FABQ subscales have shown significant association between them. (TSK and FABQ physical: $r = 0.39$–0.76, and TSK and the FABQ work: $r = 0.33$–0.56) (Crombez et al., 1999; Swinkels-Meewisse et al., 2003). Pfingsten et al in a validation study of a German version of the FABQ, in contrast to the original version with two subscales, suggested a three-subscale version dividing the "work factor" into two factors, one concentrating on work 'as a cause' of back pain and the other focusing on 'prognosis' in relation to the ability to work. They demonstrated satisfactory values for test–retest reliability and internal consistency of the three subscales (Pfingsten et al., 2000). Unfortunately, no further confirmation of the three-subscale structure has been published. An unanswered question remains whether the two or the three-subscale structure is appropriate for use in clinical practice.

An aim of this study was to assess fear-avoidance in a sample of Greek musculoskeletal patients and to compare findings cross-culturally with the German, Swiss-German, French and English FABQ versions. An additional objective was to address the issues of construct validity and responsiveness of the FABQ.

2. Study design

2.1. The original FABQ and the Greek version (FABQ-GR)

The FABQ is a 16-item self-report measure that assesses the individual's fear of pain associated with physical and occupational activities. Each item is scored from 0 to 6 with higher scores indicating increased levels of fear-avoidance beliefs. The FABQ contains the 4-item physical activity (FABQ-physical) subscale, with its score ranging from 0 to 24, and the 7-item work (FABQ-work) subscale with a 0 to 42 scoring range. The five remaining questions are used as delusive items (Waddell et al., 1993).

The adaptation of FABQ into Greek followed the guidelines published in the literature (Beaton et al., 2000; Swaine-Verdier et al., 2004). These included: the translation, the synthesis, the back-translation and the initial field testing phases. A team of a psychologist, a physiotherapist, an occupational therapist, a medical doctor, a low back pain sufferer whose native language was English and a teacher of English as a foreign language, translated the questionnaire into Greek encouraged to strive for idiomatic rather than word-for-word translation, according to the published guidelines (Swaine-Verdier et al., 2004). Cultural and vocabulary adaptations were agreed in a consensus meeting. Two bilingual professionals completed the back translation of the preliminary version, attempting conceptual equivalence, acceptability and adaptation of wording to the target population. No conceptual differences were noted between the two versions and the provisional-final questionnaire was tested. Field-testing of the provisional version included its completion by a small sub-selection of patients ($n = 13$) of the target group, by means of a one-to-one interviews in order to examine the potential distribution of responses, comprehension and to ensure linguistic, face and content validity. The findings of this preliminary field-testing indicated that the adapted version appeared to retain its equivalence to the original.

2.2. Subjects

The first 70 chronic LBP patients (pain duration > 3 months) who proceeded for physiotherapy in a private clinic and complied with the inclusion–exclusion criteria (Table 1) were asked to participate in the study. Four subjects were excluded (one due to pregnancy, two due to recent lumbar surgery in the last 6 months and one due to bone cancer), and they were replaced by the next four chronic LBP patients who proceeded for treatment. All patients agreed and written consent was obtained to participate in the study. During the first visit (assess-

Table 1
Inclusion and exclusion criteria for the LBP patients of the study

Inclusion criteria	Exclusion criteria
1. Current episode of low back pain > than 3 months	1. Spinal pain due to acute trauma, infection, systemic diseases
2. Age from 18 to 75	2. Presence of malignancies
3. Ability to understand written and spoken Greek	3. Spinal surgery
	4. Pregnancy

ment) ($t1$), the FABQ-GR and a battery of questionnaires were administered (see Instruments). A subgroup of patients ($N = 25$), randomly selected, was asked to complete the FABQ-GR again after 48 h ($t2$) before initiating any treatment. A further randomly selected group of patients ($N = 22$) was asked to complete for a third time the FABQ-GR ($t3$), together with their opinion if they were improved or not, after completion of a treatment protocol (see Testing the scale). The characteristics of the subjects of the study are depicted in Table 2. All parts of the study were developed within the principles and standards of the *Declaration of Helsinki* and in accordance with the *Guidelines on the Practice of Ethics Committees in Medical Research Involving Human Subjects*.

2.3. Instruments

The patients completed the following questionnaires (at $t1$):

- A general socio-demographic questionnaire in order to extract epidemiology data.
- The Greek version of the fear-avoidance beliefs questionnaire (FABQ-GR), as formulated by the adapting procedure.
- The Greek version of the Tampa Scale of Kinesiophobia (TSK-GR) (Kori et al., 1990). This is a 17-item questionnaire with score ranging from 17 to 68. Scores of 37 or less are suggestive of low fear of movement and scores of greater than 37 indicate high fear of movement. The Greek version has shown adequate validity and reliability (Georgoudis et al., 2005). The total scale score was used (Vlaeyen and Linton, 2000).
- The Greek version of the hospital anxiety and depression scale (HADs) (HAD-GR) (Zigmond and Snaith, 1983). A 14-item questionnaire (scaled 0–3) of two subscales, the 7-item HAD-Anxiety and the 7-item HAD-Depression, used to assess the levels of anxiety and depression, with validity and reliability shown for the Greek version (Georgoudis and Oldham, 2001).
- The Greek version of the pain locus of control questionnaire (PLC-GR) (Main and Waddell, 1991). A 20-item questionnaire measuring whether patients perceive that their pain can be effectively controlled by themselves or externally. It consists of two subscales: a pain control subscale (PLC-PC) that examines patients' beliefs about being able to affect their pain levels and a pain responsibility subscale (PLC-PR) that examines the extent to which patients believe that managing pain should be the physician's responsibility or something for which they have to take a degree of responsibility. The Greek version retains its equivalence to the original (Georgoudis et al., 2005 – submitted).
- The Greek version of the modified somatic perception questionnaire (MSPQ-GR) (Main and Waddell, 1991). This is a 13-item scale designed to measure psychological distress and specifically heightened somatic awareness among chronic pain patients, with a possible score from 0 to 39. The Greek version has shown adequate validity and reliability (Katsoulakis et al., 2004).
- The Visual Analogue Scale (VAS) in order to describe the average intensity of pain during the last week. This 10 cm line anchored with the phrases "no pain" and "worst possible pain" is a well-validated measure in chronic pain (Ogon et al., 1996).

Table 2
Demographic characteristics for the LBP patients participating in the study

	Demographic characteristics of the subjects		
	Sample ($t1$) $N = 70$	Sample ($t2$) $N = 25$	Sample ($t3$) $N = 22$
Men	12 (17.1%)	5 (20%)	4 (18.2%)
Women	58 (82.9%)	20 (80%)	18 (81.8%)
Age (range)	42.2 ± 12.0 (18–72)	40.3 ± 13.6 (18–72)	43.2 ± 12.2 (24–72)
General health			
Excellent	4.2%	3.8%	9.1%
Very good	27.1%	19.2%	27.3%
Good	33.3%	34.6%	27.3%
Not bad	31.3%	34.6%	36.4%
Poor	4.2%	3.8%	0%
Job status			
Office	34.7%	26.9%	27.3%
Manual (light)	45.6%	46.2%	59.1%
Manual (heavy)	19.7%	23.1%	13.6%

2.4. Procedure

Participants completed (at $t1$) the questionnaire set in a random order so to avoid bias (e.g., favouring of the first questionnaire tested). The administrator used a standardized script to explain the requirements of the questionnaires, and any questions were answered.

2.5. Testing the scale

Short-term test–retest reliability was estimated on a subgroup of 25 chronic LBP patients randomly selected from the initial sample. The questionnaire was administered to the patients for the first time ($t1$) during their initial visit to the clinic. A repeat administration ($t2$) after 48 h and before first treatment session (without any active treatment in-between) was chosen in order to minimise clinical or cognitive changes but also to reduce any chance recall of previous answers.

Responsiveness was examined for the FABQ-GR after the implementation of a behaviorally oriented physical therapy program, in a subgroup of 22 subjects ($t3$). The physiotherapy approach included traditional approaches (electrotherapy, deep friction massage, myofascial release techniques), and a strong element of behavioural intervention (patient to play an active role in his recovery, to view LBP as a common condition and not a serious disease, graded exercise), administered by the first author (PT with specialisation in pain management).

Construct validity was assessed in the form of convergent and divergent validity, together with exploratory factor analysis. Convergent (criterion-related) validity was examined by the degree of correlation between the TSK-GR and the FABQ-GR scores since both measures are developed to measure the same (fear-avoidance) construct. Divergent validity was studied by correlating the total subscale scores with variables assessing different concepts than fear-avoidance and belief; average *pain intensity* during last week assessed by a visual analogue scale (VAS: 0–10 cm); *anxiety and depression* assessed by the Greek version of the HAD; *psychological distress* assessed by the Greek version of the MSPQ; *pain control* and *pain responsibility* assessed by the Greek version of the PLC. An a priori lack of association was hypothesised between the FABQ-GR and the PLC-GR subscales since no previous data exist and these two cognitive assessment tools are constructed to measure completely different cognitive parameters of pain (discriminant construct validity).

2.6. Statistical analysis

All data inserted in the statistical analyses were examined for approximation of normal distribution (Kolmogorov–Smirnov goodness of fit test), skewness and kurtosis. Descriptive statistics and frequencies were also computed. Significance was set at $p < 0.05$ and it was adjusted when needed (Bonferroni correction). SPSS© 13.0 was employed in the analyses.

Test–retest reliability was mainly examined with the intraclass correlation coefficient (ICC) for the scores taken with 48 h difference ($t1$ and $t2$), allowing for the level of chance agreement and distribution effects. Similarly to the original paper, simple per cent agreement of retest answers was calculated and the overall coefficient of reproducibility was computed. Both individual items and subscales scores were examined using ICC.

The non-parametric Spearman rank correlation coefficient (rho – ρ) was used in all correlation analyses because a normal distribution could not be demonstrated for all the parameters studied. Spearman coefficient values were interpreted as being an excellent relationship, $\rho > 0.91$; good, 0.90–0.71; moderate, 0.70–0.51; fair, 0.50–0.31; and little or no relationship, $\rho < 0.30$ (Atkinson and Nevill, 1997). Internal consistency of the FABQ-GR was assessed with the Cronbach's alpha statistic (alpha – α), independently for each subscale and for all items together.

Much discussion exists concerning the calculation of responsiveness (Husted et al., 2000; Terwee et al., 2003). According to Terwee et al. (2003), responsiveness can be classified into three categories: (A) responsiveness as the ability to detect change in general (sensitivity to change); (B) responsiveness as the ability to detect clinically important change, and (C) responsiveness as the ability to detect real change in the concept being measured. In this study, it was attempted to calculate measures from all three categories. Specifically, it was computed:

- *From category A*. The *effect size* [$ES = \text{mean}(t1 - t3)_{\text{total group}}/SDt1_{\text{total group}}$] and the *paired t test in all patients who underwent treatment* [*p*-value]. A higher ES indicates greater sensitivity to change.
- *From category B*. The *standardised effect size* [$SES = \text{mean}(t1 - t3)_{\text{improved}}/SDt1_{\text{improved}}$] and the *paired t tests in patients who did and did not improve* [*p*-value] with the determination of important change according to the patient. A higher SES indicates greater responsiveness.
- *From category C*. The *receiver operating curve* (ROC) with determination of important change according to the patient [area under the curve –AUC]. The patients determined their improvement using a dichotomous variable (Yes or No improvement).

Factor structure was evaluated using exploratory factor analysis – principal component analysis (PCA) with varimax rotation. The varimax orthogonal rotation was selected in order to maximize the dispersion of loadings within factors (Tabachnick and Fidell, 1996). The number of factors extracted in the PCA analysis was not predetermined and only factors with eigenvalue >1 ("Kaiser–Guttman criterion") were considered to significantly contribute in explaining the variance. Their significance was further confirmed with the scree plot. Items were accepted on the final factors if they had a loading of more than 0.45 on the corresponding factor, similarly to the original study.

3. Results

3.1. Factor analysis

Due to the excessive skewness and kurtosis, the items FABQ8 and FABQ16 were removed from further analysis, as for to the original paper. The intercorrelations among the FABQ items ranged from $r = -0.38$ to $r = 0.905$. Although the degree of correlation between the items FABQ13 and FABQ14 ($r = 0.905$) was very close to the criterion of redundancy (correlation > 0.9) (Tabachnick and Fidell, 1996), these items were kept in the analysis. A separate analysis with items 13 and 14 removed did not yield better results than the former analysis, thus only one analysis is presented (with items 13 and 14 included). The singularity and multi-collinearity requirements were confirmed with the Determinant of the *R*-matrix (correlation matrix) (>0.00001). An adequate sample and distinct clusters of variables were indicated from the Kaiser–Meyer–Olkin measure of sampling adequacy (KMO = 0.75, it should be >0.5) and the Bartlett's test of sphericity ($\chi^2 = 649.1$, $df = 120$, $p < 0.001$) (Tabachnick and Fidell, 1996).

When the FABQ items 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14 and 15, were inserted in the PCA model with varimax rotation, the model gave 3 factors with eigenvalue > 1 (65% variance explained) (Table 3), which were confirmed by the scree plot. The item FABQ1 loaded similarly to factors 2 (eigenvalue 0.416) and 3 (eigenvalue 0.465) (inconsistent loading) and had low communality (0.4), indicating that it should be removed, according to the methodology followed in the original paper (Waddell et al., 1993). The 3 factors identified in this paper were similar to the German version (Pfingsten et al., 2000). In Table 3, the two versions are comparatively presented. The first factor (FABQ work1: items 12, 13, 14, 15) explained 37.5% of the variance and con-

Table 3
Factor analysis and descriptive statistics for the Greek and German (in parentheses) versions of the FABQ

Comparison between the Greek and German versions of the FABQ (3-factors)								
(German version) – Greek version								
Factor FABQ work as cause								
German: variance explained 43.4%	'*Work as a cause*'						German: $a = 0.89$	
Greek: variance explained 15.2% – eigenvalue 2.13							Greek: $a = 0.86$	
Items	Loading		Mean		SD		Median	
6. … 'caused by work'	(0.72)	0.82	(3.28)	2.83	(2.28)	2.33	(3)	3
7. … 'work aggravated pain'	(0.86)	0.78	(4.26)	3.65	(2.02)	2.20	(4)	4
9. … 'work too heavy'	(0.68)	0.49	(2.43)	1.46	(2.18)	2.01	(3)	0
10. … 'work makes pain worse'	(0.84)	0.84	(3.49)	3.30	(2.05)	1.98	(3)	3
11. … 'work might harm my back'	(0.80)	0.73	(3.38)	3.34	(2.04)	2.09	(3)	4
Factor FABQ work as prognosis								
German: variance explained 11.8%	'*Prognosis work*'						German: $a = 0.94$	
Greek: variance explained 37.5% – eigenvalue 5.24							Greek: $a = 0.90$	
Items	Loading		Mean		SD		Median	
12. … 'should not do my work'	(0.63)	0.85	(3.23)	2.51	(2.28)	2.34	(3)	3
13. … 'cannot do my normal work'	(0.70)	0.91	(3.06)	2.18	(2.32)	2.26	(3)	1
14. … 'wait until pain is treated'	(0.75)	0.92	(2.96)	1.91	(2.43)	2.23	(3)	1
15. … 'no return within 3 months'	(0.80)	0.68	(2.04)	0.97	(2.19)	1.63	(2)	0
16. … 'no return at all'	(0.59)	[a]	(1.71)	0.84	(2.15)	1.71	(1)	0
Factor FABQ physical								
German: variance explained 8.9%	'*Physical activity*'						German: $a = 0.69$	
Greek: variance explained 12.3% – eigenvalue 1.72							Greek: $a = 0.72$	
Items	Loading		Mean		SD		Median	
1. … 'caused by physical activity'	(0.59)	0.47[b]	(3.86)	3.13	(1.95)	1.99	(4)	3
2. … 'physical activity worsens pain'	(0.67)	0.55	(4.47)	3.59	(1.74)	2.00	(5)	4
3. … 'physical activity might harm'	(0.68)	0.81	(3.45)	3.24	(2.33)	1.97	(3)	3
4. … 'better no physical activity'	(0.69)	0.77	(4.45)	4.37	(1.78)	1.97	(5)	5
5. … 'cannot do physical activity'	(0.59)	0.63	(3.75)	3.36	(1.89)	2.21	(3)	3

German: total variance explained 64.1% – Greek: total variance explained 65%.

[a] Removed before the analysis due to extreme skewness and kurtosis.

[b] Also loaded to Factor 2 (0.42) (inconsistent loading) – removed from further analysis.

sisted of items which relate to patients' assumptions about their probable return to work ('work prognosis'). The second factor (FABQ work2: items 6, 7, 9, 10, 11) accounted for 15.2% of the variance explained. The items represent patients' beliefs about characteristics of the work environment which could be responsible for their current low back pain ('work as a cause'). The single 'work factor' of the original version of the FABQ was in this way split into two separate factors. The third factor (FABQ physical: items 2, 3, 4, 5) was identical to the second factor of the English version ('physical activity') and accounted for 12.3% of the total variance (Table 3).

Concerns about the orthogonality of the subscales have been raised due to the consistently found correlation between them (Crombez et al., 1999; Staerkle et al., 2004; Swinkels-Meewisse et al., 2003; Waddell et al., 1993). Therefore a non-orthogonal, the direct oblimin rotation method which permits correlation among factors, was also applied (delta set at 0). Not any different results were obtained from applying the oblimin rotation method.

3.2. Internal consistency

Cronbach's α for the FABQ physical (2, 3, 4, 5) was $\alpha = 0.72$, for the FABQ work1 prognosis (12, 13, 14, 15) was $\alpha = 0.90$, and for the FABQ work2 as cause (6, 7, 9, 10, 11) was $\alpha = 0.86$.

3.3. Test–retest

ICC values for the 16 FABQ items ranged from 0.69 to 0.97 with a mean ICC of 0.86, indicating a very good reproducibility for the individual items. ICC values for the 3-factor structure as identified in this study were: FABQ physical (2, 3, 4, 5): 0.85 (95% CI: 0.65–0.93), FABQ work1 prognosis (12, 13, 14, 15): 0.93 (95% CI: 0.84-0.97) and FABQ work2 as cause (6, 7, 9, 10, 11): 0.94 (95% CI: 0.87–0.97). Identical answers were obtained in 241 out of 400 answers, giving a reproducibility coefficient of 60.25%.

3.4. Responsiveness

3.4.1. Responsiveness as sensitivity to change

Statistically significant reductions pre- and post-treatment were noted for all subscales (paired t tests): FABQ physical pre- vs. post-Rx ($t = 3.074$, $p < 0.01$); FABQ work1 (6,7,9,10,11) pre- vs. post-Rx ($t = 4.899$, $p < 0.001$); and FABQ work2 (12,13,14,15) pre- vs. post-Rx ($t = 2.175$, $p < 0.05$) (Table 4). The ES calculated were for the FABQ work1: ES = 0.38, FABQ work2: ES = 0.30, and FABQ physical: ES = 0.52 (Table 4).

3.4.2. Responsiveness as the ability to detect clinically important change

Statistically significant reductions pre- and post-treatment were noted in the patients who did improve (the important change determined by the patient) ($p < 0.02$) (Table 5), whereas non-significant changes were noted in patients who did not improve ($p > 0.3$) (Table 5). The SES was for the FABQ work1 = 0.35, the FABQ work2 = 0.53 and the FABQ_Phys = 0.55 (Table 5).

3.4.3. Responsiveness as the ability to detect real change in the concept being measured

The ROC (with the important change determined by the patient) produced for the FABQ physical: AUC = 0.682, FABQ work1: AUC = 0.665, FABQ work2: AUC = 0.600 (Fig. 1).

3.5. Convergent construct validity

As expected, both FABQ work and FABQ physical subscales significantly correlated with TSK (FABQ physical: $r = 0.55$, $p < 0.001$, FABQ work1: $r = 0.39$, $p < 0.001$, FABQ work2: $r = 0.25$, $p < 0.05$). Interestingly, the correlations were only of fair to moderate magnitude.

3.6. Divergent construct validity

Little to fair relationship was noted between FABQ physical, MSPQ ($\rho = 0.20$, p=0.053), PLC pain responsi-

Table 4
Responsiveness as the ability to detect change in general (sensitivity to change) for the FABQ subscales

	Paired differences								
	Mean	*N*	SD	Mean	SD	*t*	df	sig	ES
FABQ work1 pre-Rx	4.3	20	5.4	2.90	5.96	2.175	19	<0.05	0.38
FABQ work1 post-Rx	7.2	20	7.6						
FABQ work2 pre-Rx	11.8	20	8.5	2.75	2.51	4.899	19	<0.001	0.30
FABQ work2 post-Rx	14.6	20	9.1						
FABQ physical pre-Rx	13.8	22	6.1	2.36	3.61	3.074	21	<0.01	0.52
FABQ physical post-Rx	11.4	22	5.5						

Paired t test and effect sizes (ES) were calculated for the total group of patients receiving treatment.

Table 5
Responsiveness as the ability to detect clinically important change for the FABQ subscales

				Paired differences					
	Mean	N	SD	Mean	SD	t	df	sig	SES
No improvement									
FABQ work1 pre-Rx	17.0	5	4.5	−1.60	3.21	0.638	4	0.558	N/A
FABQ work1 post-Rx	15.4	5	9.2						
FABQ work2 pre-Rx	3.6	5	4.5	1.20	4.21	−1.115	4	0.327	N/A
FABQ work2 post-Rx	4.8	5	4.3						
FABQ physical pre-Rx	13.6	5	7.8	0.60	1.82	0.739	4	0.501	N/A
FABQ physical post-Rx	14.2	5	7.1						
Improved									
FABQ work1 pre-Rx	13.7	15	8.8	−3.13	2.23	−5.437	14	<0.001	0.35
FABQ work1 post-Rx	10.6	15	8.2						
FABQ work2 pre-Rx	8.4	15	8.1	−4.27	5.93	−2.785	14	<0.02	0.53
FABQ work2 post-Rx	4.1	15	5.9						
FABQ physical pre-Rx	13.8	17	5.8	−3.24	3.56	−3.744	16	<0.002	0.55
FABQ physical post-Rx	10.6	17	4.9						

Paired t test and standardised effect sizes (SES) were calculated for the group of patients that improved after receiving treatment.

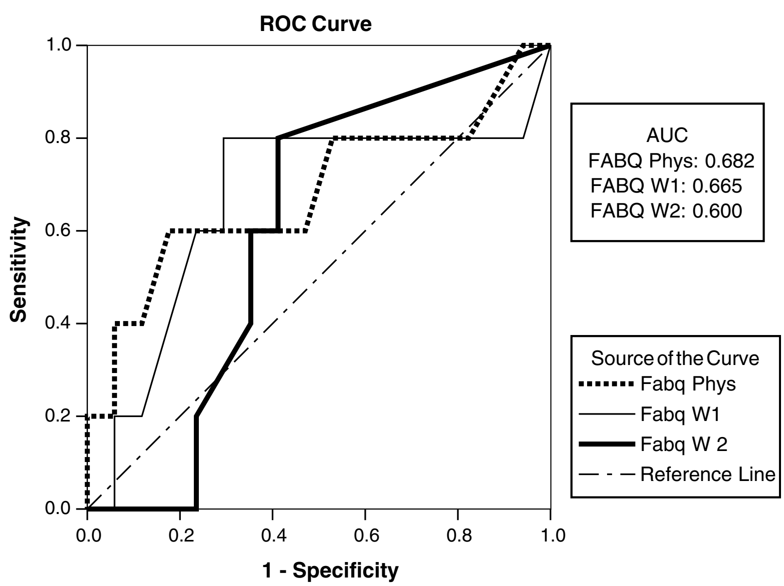

Fig. 1. Responsiveness of the FABQ-GR as the ability to detect real change in the concept being measured. The receiver operating curve (ROC) and the area under the curve (AUC) were calculated using as comparison the improved vs. the not-improved group.

bility ($\rho = -0.30$, $p < 0.01$), HAD-depression ($\rho = 0.29$, $p < 0.01$) and VAS ($\rho = 0.49$, $p < 0.001$) (Table 6). The FABQ work1 prognosis (12, 13, 14, 15) subscale showed moderate relationship with the VAS ($\rho = 0.56$, $p < 0.001$), fair with the HAD-depression ($\rho = 0.39$, $p < 0.01$), PLC-pain responsibility ($\rho = -0.30$, $p < 0.01$), and little with HAD-anxiety ($\rho = 0.26$, $p < 0.05$). The FABQ work2 as a cause (6, 7, 9, 10, 11) subscale showed fair relationship with the HAD-depression ($\rho = 0.35$, $p < 0.01$) and little with HAD-anxiety subscales ($\rho = 0.22$, $p < 0.05$) (Table 6). It was interesting that the FABQ work2 subscale, showed no association with the intensity of pain (VAS) ($p > 0.1$) (Table 6). The association among the three subscales was: FABQ physical/work1: $\rho = 0.31$, $p < 0.005$; FABQ physical/work2: $\rho = 0.41$, $p < 0.001$; FABQ work1/work2: $\rho = 0.40$, $p < 0.001$.

4. Discussion

Among the study aims was the desire to assess fear-avoidance in Greek musculoskeletal patients with a

Table 6
Bivariate correlations of the FABQ subscales with MSPQ, PLC, HAD, TSK and VAS

Divergent construct validity of the FABQ			
($N = 70$) Spearman's rho (ρ)	FABQ work2	FABQ work1	FABQ physical
MSPQ	0.18	0.02	0.20
Pain control (PLC-PC)	−0.05	0.10	0.12
Pain responsibility (PLC-PR)	−0.15	−0.30[b]	−0.30[b]
HAD-anxiety	0.22[a]	0.26[a]	0.10
HAD-depression	0.35[b]	0.39[b]	0.29[b]
Pain Intensity (VAS) ($N = 40$)	0.12	0.56[c]	0.49[c]

FABQ, fear-avoidance beliefs questionnaire; MSPQ, modified somatic perception questionnaire; PLC, pain locus of control; HAD, hospital anxiety and depression scale; VAS, visual analogue scale.
All other correlations are non-significant.
[a] Correlation is significant at the 0.05 level (2-tailed).
[b] Correlation is significant at the 0.01 level (2-tailed).
[c] Correlation is significant at the 0.001 level (2-tailed).

newly validated FABQ version (FABQ-GR). Using cross-cultural comparisons with existing versions of FABQ, the findings indicate that the FABQ-GR is a valid, reliable and responsive version, which retains the psychometric properties of the original. Other objectives were to investigate the construct validity and structure of the FABQ. The results confirmed, using a correlational design, new aspects of convergent and divergent validity, indicating further clinical value and significance for the measure. A factor analysis provided a 3-factor model which was shown to adequately respond to the validity and reliability requirements.

4.1. Factor analysis

The exploratory factor analysis was used to examine the structure of the FABQ-GR instead of a confirmatory factor analysis model, since the number of possible factors expected was not predetermined from the literature and either a two or a three factor model was anticipated (Pfingsten et al., 2000; Waddell et al., 1993).

PCA modelling identified three distinct factors with salient loadings of the items. The FABQ physical subscale was identical to the 2-factor structure quoted in previous studies (Chaory et al., 2004; Pfingsten et al., 2000; Staerkle et al., 2004; Waddell et al., 1993), confirming its unambiguous validity. The FABQ 'work' subscales were similar to the German version (Pfingsten et al., 2000) (Table 3), indicating statistical soundness for these subscales, as well. The exclusion of the FABQ1 item from the FABQ physical subscale, in opposition to the Pfingsten and co-workers' (2000) view did not seem to change the findings, partially confirming Waddell's et al. (1993) suggestion for exclusion.

A serious concern in factorial models is the adequacy of sampling, resulting in desired samples consisting of 300 and more subjects (Tabachnick and Fidell, 1996). Although, in this study only 70 subjects participated, the factors identified had more than four loadings above eigenvalues >0.6 (Table 3), confirming a reliable model regardless of sample size (Guadagnoli and Velicer, 1988). Therefore, it can be argued that the 3-factor model, as established in this study, is statistically sound and acceptable for use.

4.2. Internal consistency

Comparable results were produced for the internal consistency with the German version (Table 3). This is indicative of the homogeneity of the items which form acceptable subscale measures. The internal consistency of the physical subscale ($\alpha = 0.72$) was above the critical level of 0.70 (Streiner and Norman, 1995), in contrast to the German version ($\alpha = 0.69$). This value was not however the highest quoted in the literature (0.52–0.83), though very close to the original version (0.77) (Waddell et al., 1993). Similarly close to the original ($a = 0.88$), were the values found for the FABQ work1 and work2 subscales ($\alpha = 0.90$ and $\alpha = 0.86$), despite the wide range of values reported in the literature ($\alpha = 0.68$–0.92). The common characteristics of the patients – chronic spinal pain patients – participating in the original and the present study may be responsible for the close findings.

4.3. Test–retest

The coefficient of reproducibility (60.25%) and the ICC (0.86) calculated for the 16 individual items of the FABQ-GR are very similar to the literature (Staerkle et al., 2004: 61% and 0.76; Waddell et al., 1993: 71% and 0.74). Furthermore, all subscales showed excellent ICC values (0.85–0.94), among the highest reported (0.64–0.95).

4.4. Responsiveness

The FABQ physical activity was shown to be the most responsive subscale in all categories of responsiveness (highest AUC, SES and ES). The FABQ work1 was rated as second most responsive subscale according to the AUC and the ES, and the FABQ work2 followed. This finding corroborates previous data where the FABQ 'physical' was shown to be more responsive than the FABQ 'work' (0.89 and 0.27) (Woby et al., 2004). The present results regarding responsiveness may be of essential clinical value, especially when using FABQ as an outcome measure in research designs. In this way, the research question posed by Waddell for the responsiveness to treatment of the FABQ subscales seems to have started to be addressed (Waddell et al., 1993).

4.5. Construct validity

The FABQ subscales showed fair inter-correlation among them (0.31–0.41). The magnitude of association for the three subscales indicates three separate clusters of variables which are only fairly correlated. This would be expected when independent parameters of the same construct are measured (Streiner and Norman, 1995). The moderate to high intercorrelations among scales, especially between the work subscales ($r = 0.67$–0.71), showed by the German and Swiss-German versions (Pfingsten et al., 2000; Staerkle et al., 2004), could not satisfactorily confirm the independence of the two 'work' subscales. This study's findings for the first time explicitly validate the existence of the two separate 'work' subscales and confirm a 3-factor model for the FABQ.

The convergent validity of the FABQ was supported by the pattern of correlations with the TSK. All subscales showed from little to moderate association with the TSK score confirming previous studies (Crombez et al., 1999; Swinkels-Meewisse et al., 2003) and the scale's criterion-related validity. The fact that the calculated correlation between FABQ and TSK does not exceed 'moderate' although both scales are considered to measure fear-avoidance, might mean that the theoretical constructs are not completely the same (although they show an empirical 'moderate' overlap) (Swinkels-Meewisse et al., 2003).

The validity of the FABQ was reinforced by its divergent construct assessment. The independent construct of the pain locus of control (PLC-pain control), although considered to be a parameter of pain cognitive assessment, it is not associated with fear avoidance. This new finding in the literature of no association between the two constructs differentiates pain-control from fear-avoidance beliefs. Another interesting finding with clinical value was the little but significant inverse association with the PLC-pain responsibility subscale ($r = -0.21$ to $r = -0.30$). This may indicate that when in fear and in avoidance of activities, the patients' pain responsibility is diminished. The specific finding confirms previous statement that in fearful patients voluntary recruitment of coping strategies might be hampered (Crombez et al., 1999).

The lack of association of fear avoidance and psychological distress (MSPQ) reported in this study is in contradistinction with previous research where fair though significant associations were shown (Table 7). Although, a trend towards this direction may be shown from the data ($p < 0.1$) (Table 6), it is difficult to explain this lack of association in this study. Possibly the relatively small sample size did not allow for statistical significance.

In agreement with the literature (Staerkle et al., 2004; Waddell et al., 1993), depressive symptoms were fairly associated with all FABQ subscales, and anxiety with the FABQ work subscales. This is a consistent finding (Tables 6 and 7) verifying the frequent situation where phobic patients feel also anxious and depressed, especially when these theoretically independent constructs, in clinical terms, are not that independent and better seem to describe the general psychological status of the patient (Vlaeyen et al., 1995).

In contrast to some previous research findings (Pfingsten et al., 2000; Waddell et al., 1993) and in agreement with others (Staerkle et al., 2004), the FABQ physical factor was largely dependent on the pain intensity measure (VAS) (Table 7). This finding seems to actually confirm the view that fear-avoidance measures are capable of predicting the intensity of chronic musculoskeletal pain (Vlaeyen and Linton, 2000). In order to compare our data with the literature, the two work subscales were combined in a single FABQ work subscale, and the correlations were re-calculated (Table 7). The single FABQ work subscale and the VAS produced low magnitude correlation coefficients indicative of little or no relationship. This is in agreement with all previous relevant data which primarily assessed the validity and reliability of the German, French and English versions (Chaory et al., 2004; Pfingsten et al., 2000; Waddell et al., 1993). The only disagreement comes from a study which assessed the validity of the Swiss-German version (Staerkle et al., 2004) (Table 7).

Table 7
Relationship of FABQ with MSPQ, PLC, HAD and VAS (literature and present study's findings)

	FABQ work[d]		FABQ physical	
	This study[d]	Literature	This study	Literature
MSPQ	0.18	0.36[c] (1) 0.36[c] (4)	0.20	0.31–0.36[c] (1) 0.19 (4)
HAD-anxiety	0.27[a]	0.30[b] (2) 0.20 (3)	0.10	0.14 (2) 0.17 (3)
HAD-depression	0.44[b]	0.42–0.45[c] (1) 0.29[a] (2) 0.13–0.18 (3) 0.41[c] (4)	0.29[b]	0.32–0.37[c] (1) 0.33[a] (2) 0.15 (3) 0.36[c] (4)
Pain intensity (VAS)	0.28[a]	0.47–0.60[c] (1) 0.28[b] (2) 0.26–0.38 (3) 0.23[b] (4)	0.49[c]	0.48[c] (1) 0.24[b] (2) 0.24 (3) 0.12 (4)

(1) Staerkle et al. (2004); (2) Chaory et al. (2004); (3) Pfingsten et al. (2000); (4) Waddell et al. (1993).
FABQ, fear-avoidance beliefs questionnaire; MSPQ, modified somatic perception questionnaire; PLC, pain locus of control; HAD, hospital anxiety and depression scale; VAS, visual analogue scale.
All other correlations are non-significant.
[a] Correlation is significant at the 0.05 level (2-tailed).
[b] Correlation is significant at the 0.01 level (2-tailed).
[c] Correlation is significant at the 0.001 level (2-tailed).
[d] In order to compare findings in the literature, the FABQ work1 and work2 subscales were combined in one FABQ work subscale. All correlations were re-calculated for the single FABQ work subscale.

4.6. Limitations

Although much discussion exists concerning the definition and calculation of responsiveness (Husted et al., 2000; Terwee et al., 2003), it is accepted that the assessment of the responsiveness of an instrument should be performed in evaluating the occurred change within the context of a randomised control trial (RCT). In this study there was only one group of patients receiving a behaviourally oriented physiotherapy treatment. There was no RCT design and the number of treated subjects was only 22 (17 improved according to their rating). Therefore, the measures of responsiveness calculated should be interpreted with caution and further confirmed in future studies with an appropriate RCT design.

It is necessary to point out that the sample of the study is relatively low ($N = 70$) (with a gender bias – females = 58). This should be a reason for some caution in interpreting the results. It should also be emphasized that this study was designed cross-sectionally and any significant correlations should not be confused with causal effects. Another issue of concern regarding this study is that only self-reported measures were used. The analysis should have included functional indices and/or other behavioural performance tests, so to further secure external validity for the FABQ-GR. However, such associations have been explicitly examined for other adapted versions of the FABQ (Chaory et al., 2004; Pfingsten et al., 2000).

5. Conclusions

New aspects of construct validity and an alternative factorial model for the FABQ with three subscales were identified, further advancing the understanding and clinical applicability of this fear-avoidance measure. The satisfactory validity, reliability and responsiveness of the 3-factor FABQ-GR shown in this study, make it suitable for clinical use with Greek musculoskeletal pain patients.

Acknowledgements

We acknowledge Prof. Chris Main's, clinical psychologist, useful comments in the initial planning of the study, Prof. Jacqueline Oldham for advice, reviewing and proofreading the final version, and the time and effort of the anonymous reviewers of the European Journal of Pain spend on reviewing previous versions of our manuscript. Also, we acknowledge the assistance of Ms. Marmarinou Aleksandra, PT, in data collection.

References

Atkinson G, Nevill A. Comment on use of concordance correlation to assess the agreement between two variables. Biometrics 1997;52: 775–7.

Beaton D, Bombardier C, Guillemin F. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000;25:3186–91.

Chaory K, Fayad F, Rannou F, Lefevre-Colau M, Fermanian J, Revel M, et al. Validation of the French version of the fear avoidance belief questionnaire. Spine 2004;29(8):908–13.

Crombez G, Vlaeyen J, Heuts P, Lysens R. Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. Pain 1999;80:329–39.

Georgoudis G, Papathanasiou G, Spiropoulos P, Katsoulakis K. Physiotherapy assessment in painful musculoskeletal conditions: the validation of the Greek version of the Tampa Scale of Kinesiophobia (pilot study). In: 1st international congress on "Pain meets general practice" organized by the World Institute of Pain and the European Federation of IASP Chapters, Sofia, Bulgaria; 2005. p. 214 [Abstracts Book].

Georgoudis G, Oldham JA. Anxiety and depression as confounding factors in cross-cultural pain research studies: validity and reliability of a Greek version of the hospital anxiety and depression scale. Physiotherapy 2001;87(2):92–3.

Guadagnoli E, Velicer W. Relation of sample size to the stability of component patterns. Psychol Bull 1988;103:265–75.

Husted J, Cook R, Farewell V, Gladman D. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000;53:459–68.

Katsoulakis C, Georgoudis G, Ntavatzikos A. Assessment of cognitive parameters of pain: the Greek version of the Modified Somatic Perception Questionnaire [Αξιολόγηση των γνωσηακών παραμ έτρων του πόνου: η Ελληνηκή έκδοση του Modified Somatic Perception Questionnaire (MSPQ)]. In: 18th Hellenic Physiotherapy Congress organized by the Hellenic Scientific Society of Physiotherapy, Athens; 2004. p. 58–9 [Abstract Book].

Kori S, Miller R, Todd D. Kinesiophobia: a new view of chronic pain behaviour. Pain Manage 1990:35–43.

Lethem J, Slade P, Troup J, Bentley G. Outline of a fear-avoidance model of exaggerated pain perception. Behav Res Ther 1983;21: 401–8.

Linton S. A review of psychological risk factors in back and neck pain. Spine 2000;25(9):1148–56.

Main C, Waddell G. A comparison of cognitive measures in low back pain: statistical structure and clinical validity at initial assessment. Pain 1991;46:287–98.

Ogon M, Krismer M, Sollner W, Kantner-Rumplmair W, Lampe A. Chronic low back pain measurement with visual analogue scales in different settings. Pain 1996;64(3):425–8.

Pfingsten M, Kroner-Herwig B, Leibing E, Kronshage U, Hildebrandt J. Validation of the German Version of the fear-avoidance beliefs questionnaire (FABQ). Eur J Pain 2000;4:259–66.

Staerkle R, Mannion A, Elfering A, Junge A, Semmer N, Jacobshagen N, et al. Longitudinal validation of the fear-avoidance beliefs questionnaire (FABQ) in a Swiss-German sample of low back pain patients. Eur Spine J 2004;13:332–40.

Streiner DL, Norman GR. Health measurement scales. 2nd ed. Oxford: Oxford University Press; 1995.

Swaine-Verdier A, Downward L, Hagell P, Thorsen H, McKenna S. Adapting quality of life instruments. Value Health 2004;7(Suppl. 1): S27–30.

Swinkels-Meewisse E, Swinkels R, Verbeek A, Vlaeyen J, Oostendorp R. Psychometric properties of the Tampa Scale for Kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. Man Ther 2003;8:29–36.

Tabachnick B, Fidell L. Using multivariate statistics. 3rd ed. New York: Harper & Row; 1996.
Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res 2003;12:349–62.
Vlaeyen J, Kole-Snijders A, Boeren R, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. Pain 1995;62:363–72.
Vlaeyen J, Linton S. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain 2000;85:317–32.
Waddell G, Newton M, Henderson I, Somerville D, Main C. A fear avoidance beliefs questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain 1993;52:157–68.
Woby S, Watson P, Roach N, Urmston M. Are changes in fear-avoidance beliefs, catastrophizing, and appraisals of control, predictive of changes in chronic low back pain and disability? Eur J Pain 2004;8:201–10.
Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361–70.