

The McGill Pain Questionnaire, Japanese version, reconsidered: Confirming the theoretical structure

Mamoru Hasegawa MD PhD¹, Suguru Hattori MA², Masaru Mishima MD³, Isao Matsumoto MD PhD³,
 Takashi Kimura MD¹, Yasuko Baba MD¹, Osami Takano MD¹, Toshio Sasaki MD¹,
 Kazue Kanemura MD¹, Katsuhiko Senami MD¹, Toshinari Shibata MD¹

M Hasegawa, S Hattori, M Mishima, et al.
The McGill Pain Questionnaire, Japanese version, reconsidered: Confirming the theoretical structure.
Pain Res Manage 2001;6(4):173-180.

BACKGROUND: Based on a tripartite theoretical model of pain, the Pain Rating Index (PRI) of the McGill Pain Questionnaire (MPQ) continues to be one of the most frequently used instruments to measure clinical pain. However, language and cultural barriers have hindered its wide use and standardization in Japan. Although a number of exploratory factor analysis studies have failed to support consistently the theoretical structure of the MPQ, a few previous confirmatory factor analysis studies did statistically support the a priori model.

OBJECTIVE: To test, through confirmatory factor analysis (CFA), the theoretical structure of a Japanese version of the MPQ (JMPQ), which followed a format similar to that of the original MPQ.

DESIGN: This study used CFA on prospectively collected data from 199 consecutive outpatients with chronic pain at a university hospital to test the theoretical structure of the JMPQ.

RESULTS AND CONCLUSIONS: CFA was completed on the first 16 PRI subclass scores; this process yielded a well fitting final model that explained 92% of the covariance in the observed data. The results supported the hypothesis that the sensory, affective and evaluative subscales of the PRI are representative of the

multidimensionality of the pain experience, with minimal overlap. It is suggested that the theoretical structure of the MPQ is maintained in the JMPQ used in this study. Therefore, this study is the first step toward standardization of the JMPQ, serving as a cultural bridge in the field of pain medicine between Japan and English-speaking nations such as Canada.

Key Words: Amos; Chronic pain; Confirmatory factor analysis; Japanese version of the McGill Pain Questionnaire; JMPQ; McGill Pain Questionnaire; MPQ

Version japonaise du questionnaire de McGill sur la douleur : confirmation de la structure théorique

CONTEXTE : L'indice de cotation de la douleur du questionnaire de McGill, qui repose sur un modèle théorique composé de trois parties, continue à être l'un des instruments de mesure les plus utilisés pour évaluer la douleur clinique. Toutefois, des barrières de langue et de culture en ont empêché l'utilisation répandue et la normalisation au Japon. Même si différentes analyses factorielles exploratoires n'ont pas toujours confirmé la structure théorique du questionnaire, certaines analyses factorielles antérieures de confirmation ont cor-

voir page suivante

¹Department of Anesthesiology, Kanagawa Cancer Center, Kanagawa, Japan; ²Department of Psychiatry, Gunma University School of Medicine, Gunma, Japan; ³Department of Anesthesiology, Saitama Medical School, Saitama, Japan

Correspondence and reprints: Dr Mamoru Hasegawa, Department of Anesthesiology, Kanagawa Prefectural Cancer Center, 1-1-2 Nakao, Asahi-ku, Yokohama, Kanagawa 241-0815, Japan. Telephone +81-45-391-5761, fax +81-45-361-4692, e-mail m-hasegawa@cam.hi-ho.ne.jp

Received for publication August 28, 2000. Accepted December 10, 2000

robéré, sur le plan statistique, la validité du modèle empirique.

OBJECTIF : Vérifier, au moyen d'une analyse factorielle de confirmation (AFC), la structure théorique de la version japonaise du questionnaire, qui ressemble à celle du questionnaire original.

PLAN D'ÉTUDE : Les auteurs ont appliqué l'AFC à des données prospectives, recueillies dans un hôpital universitaire et obtenues de 199 patients externes, consécutifs, souffrant de douleurs chroniques pour vérifier la structure théorique de la version japo-naise du questionnaire.

RÉSULTATS ET CONCLUSION : L'AFC porte sur les seize premières cotations de sous-classe de l'indice; l'étude a permis de

dégager un modèle final bien adapté, pouvant expliquer 92 % de la covariance observée dans les données recueillies. Les résultats confirment l'hypothèse selon laquelle les sous-échelles sensorielle, affective et évaluative de l'indice sont représentatives du caractère multidimensionnel de la douleur, et ce, avec un minimum de chevauchement. Aussi est-il proposé de maintenir la structure théorique du questionnaire dans sa version japonaise. L'étude constitue la première étape vers la normalisation de la version japonaise du questionnaire et sert de pont culturel dans le domaine du traitement de la douleur entre le Japon et les pays anglophones comme le Canada.

The measurement of pain, particularly chronic pain, is difficult due to its complexity. One of the most important roles of clinicians is to evaluate – precisely and objectively – the subjective experience of pain. It also has become important to use universally accepted scales to evaluate pain so that data can be compared internationally.

The McGill Pain Questionnaire (MPQ) is one of the most widely used and sensitive clinical tools for the verbal assessment of pain (1). The MPQ is theoretically grounded in the gate control theory of pain (2) and is designed to assess the three dimensions of the pain experience (sensory, affective and evaluative) hypothesized by this theory. The MPQ comprises 20 subclasses of qualitatively and quantitatively ordered pain descriptors. The three interrelated, yet distinct, dimensions of the pain experience (sensory, affective and evaluative) are derived from the first 16 subclasses of these descriptors. These, as well as the remaining four sets of descriptors, which comprise a miscellaneous dimension, are summed to yield a total Pain Rating Index (PRI) based on the rank values of words chosen. The descriptors describe sensory (subclasses 1 to 10), affective (subclasses 11 to 15), evaluative (subclass 16), and mixed sensory, affective and evaluative (subclasses 17 to 20) aspects of pain. Previous studies using factor analysis revealed that these descriptors can be classified into one of three main subscales (sensory, affective and evaluative descriptors) or a miscellaneous fourth subscale (3). Each subclass has several descriptors ranked in order of intensity (rank value) beginning at the value 1, increasing one by one in series. The PRI is the sum of the rank values selected by the patient in each subclass. The score of each subscale is briefly described as the sensory score of the PRI, the affective score of the PRI, the evaluative score of the PRI (PRI-E) and the miscellaneous score of the PRI. The total score of the PRI was the sum total of each subscale score. Its purpose is to assess not only pain intensity, but also the multiple dimensions of the pain experience. Despite widespread acceptance of the MPQ in the field, many results from studies that have tested the construct validity of this instrument have been inconsistent.

Exploratory factor analyses have not consistently supported the presence of the three nonredundant factors predicted by Melzack's postulated model. Indeed, some factor analysis procedures identified two factors (4); others identified four (5,6), five (7), six (8,9) and seven factors (10).

However, exploratory factor analyses are poorly suited to test a priori postulated factor structures. On the other hand, confirmatory factor analysis (CFA) enables the direct testing of factorial structures (11-13). Following a confirmatory methodology, Turk et al (11) and Lowe et al (12) confirmed Melzack's postulated tricomponent structure. In a recent study, Donaldson (13) reported that the a priori three-factor structure provides an acceptable fit according to data but that other models may fit as well or better. Donaldson (13) found three underlying dimensions that differ somewhat from the theoretical organization implied by the gate control theory (2); this alternative structure was called the semantic model. This model distinguishes three factors: sensory action, sensory evaluative factors and affective evaluative factors. Examination of the descriptors composing the PRI subclasses may cluster about three underlying semantic dimensions that differ somewhat from the theoretical organization implied by the gate control theory. Specifically, the temporal, spatial, punctate pressure, incisive pressure, constrictive pressure and traction pressure descriptors express semantic content of sensory action. The thermal, brightness, dullness, sensory miscellaneous, tension and evaluative descriptors connote sensory evaluation. The autonomic, fear, punishment and affective miscellaneous subclasses represent affective evaluation. (It is helpful when examining these groupings to attend to the descriptors themselves, rather than the subclass names, which are useful but not infallible summaries of their semantic content.) Generally, the distinction between the sensory and affective dimensions has held up extremely well, but there is still considerable debate about the separation of the affective and evaluative dimensions. The semantic model does not assume a separate evaluative factor.

Many investigators have reported moderately high correlations between the sensory, affective and evaluative factors (3,5,6,11-17). Turk et al (11) found that cross-construct correlations were higher than within-construct correlations (reliability indexes). It was concluded that the subscales are, therefore, not different and that only the total score should be used.

However, differences in languages and cultural backgrounds have hindered the wide use and standardization of a Japanese version of the MPQ (JMPQ) in Japan. The MPQ has been translated and reconstructed into many languages worldwide. In each translation, alterations have been made, where necessary, to make the questionnaire most relevant to

the language. In the cross-cultural translation of psychometric tests, the 'decentring' concept of the back-translation technique is important (18). Decentring, a translation concept first outlined by Werner and Campbell (18), is the process by which one set of materials is translated into another language; it is not translated with as little change as possible but is translated to produce a smooth, natural-sounding version in the second language. Thus, using the back-translation technique, the language is translated more freely, so that the final image depicted in the translated version becomes as close as possible to that of the original. The major advantage of back-translation is that it gives researchers some control over the instrument development stage, because they can examine original and back-translated versions and make inferences about the quality of the translation. Because of such advantages, the back-translation technique was used in this study to create the JMPQ, which followed a format similar to that of the original MPQ (18). The authors previously used this JMPQ in patients with chronic pain to demonstrate the reliability and validity of the JMPQ as a pain rating scale (19). The original MPQ employed four mutually independent subscales to assess different qualitative aspects of pain. The subscales of the JMPQ did not display adequate discriminant validity compared with the original MPQ depending on the results of our previous paper (1,19). The four PRI components had relatively high intercorrelations within the JMPQ – similar to the English version (11). Therefore, the JMPQ had the same problems as the English version in assessing the various qualitative aspects of pain. These problems are likely a structural flaw of the original MPQ as opposed to structural changes caused by translation. As mentioned previously, a number of exploratory factor analysis studies have consistently failed to support the theoretical structure of the MPQ; however, a few previous CFA studies of chronic and acute pain statistically supported the a priori model (11-13,16,20).

The aim of this study was to confirm, through CFA, the theoretical structure of the JMPQ, which followed a format similar to that of the original MPQ.

PATIENTS AND METHODS

The present study evaluated 199 consecutive Japanese patients with chronic pain (96 male and 103 female; mean age 54.7 ± 14.1 years) who consulted the pain clinic at Saitama Medical School Hospital (Saitama, Japan) from January to May 2000. Three per cent had less than secondary education, 7% had some secondary education, 43% completed high school, 25% had some postsecondary education and 22% had completed postsecondary education. The average duration of pain was 15.3 ± 7.2 months. To minimize the effects of treatment, patients were selected at their initial visit to pick up patients without any previous treatment experience. Regions of pain described by patients were the lower back ($n=85$), neck ($n=58$), face and/or head ($n=24$), shoulder ($n=17$) and other ($n=15$). Informed consent was obtained from each subject before data collection.

The JMPQ and other pain rating scales, such as the visual analogue scale (VAS), the verbal rating scale (VRS) and the numerical rating scale (NRS), were administered, and history (concerning previous treatment, regions of pain, psychological state, duration and intensity of pain, past history, present illness and educational level, etc) was recorded in a private interview of patients awaiting treatment. There was no variation in the patients' understanding of the JMPQ according to educational attainment. Therefore, no patients were excluded from the study because of their inability to comprehend the questionnaire. Patients were allowed to choose less than one descriptor in each subclass. The JMPQ used in this study, which followed a format similar to that of the original MPQ, comprised 78 pain descriptors in 20 subclasses and a five-point intensity scale (the present pain intensity scale) (1). Pain descriptors adopted in this study had been translated by Satow et al (21). Back-translation procedures were completed by a bilingual English linguist to ensure the adequacy of the Japanese version (18). Back-translation means 'backward and forward translation' by bilingualists. As in the Turk et al study (11), only the first 16 subclasses of the PRI were included in the CFA. Subclasses 17 through 20 were excluded from analysis, because they were labelled as miscellaneous items and were not classified according to the theoretical conceptualization of pain as a sensory, affective or evaluative phenomenon.

To test the reliability of the JMPQ, internal consistency was evaluated for each subscale using the alpha coefficient. To test the validity of the JMPQ, concurrent validity was assessed for each subscale as follows: the JMPQ subscale and total scores for all patients were correlated with scores on three well known and established pain rating scales – the 100 mm VAS, the five-point VRS and the 0 to 100 NRS – by using the correlation matrix derived from the above measures and the JMPQ scales.

Data obtained in this study were analyzed using the program package Amos 4.0 (SPSS Inc, USA) (22). Using Amos, a CFA was conducted to test the hypothesized three-factor structure (Figure 1). For the purpose of identification, the first element of each congreneric set was fixed to 1.0.

The primary statistical goal in CFA is to examine how well the observed sample data fit the hypothesized model. The assessment of the overall fit of the model is based on multiple criteria that reflect the substantive meaningfulness of the model, statistical criteria (eg, the amount of variances and covariances jointly explained by the model) and practical criteria (eg, the percentage of covariance explained by the model).

The second stage of the analysis began upon findings of a less than adequate fit of the sample data for the initial hypothesized model. Thus, sensitivity analysis was completed in an exploratory fashion to delineate the cause of the model misfit. All analyses were based on covariance matrices.

Additionally, intercorrelations between subscales were compared with the reliability indexes of the subscales. Intercorrelations between subclasses (items) were also calculated to clarify the item correlations.

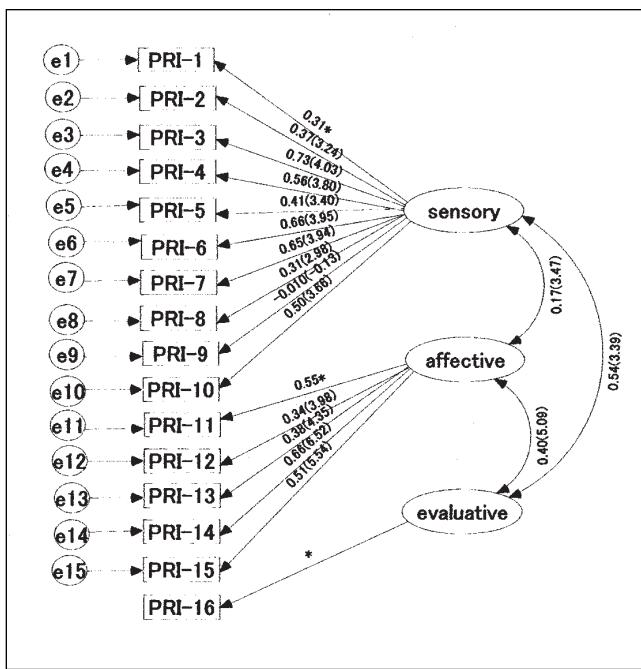


Figure 1) Standardized estimates for three-factor structure (initial model) underlying the first 16 subclasses of the Japanese version of the McGill Pain Questionnaire (MPQ). Values in parentheses are t values of the parameter estimate. Values greater than 1.96 indicate statistical significance ($P<0.05$). MPQ subscales are sensory, affective and evaluative, and MPQ subclasses are Pain Rating Index (PRI)-1 to PRI-16. Regular boxes represent observed variables and ovals represent latent (unobserved) variables; double-headed arrows represent a pattern of intercorrelation and single-headed arrows leading from the latent constructs to the boxes are regression paths that represent the link from the factors to their respective set of observed variables. *Parameter is fixed to 1.0 for purposes of statistical identification.

RESULTS

In the CFA, the 16 items of the PRI were considered separate observed variables from which the latent (sensory, affective and evaluative) variables were estimated. Table 1 presents the descriptive statistics for each of the 16 items in the sample. On the present pain intensity scale, 50 patients (25.1%) described their pain as 'mild', 62 (31.2%) as 'discomforting', 48 (24.1%) as 'distressing', 25 (12.6%) as 'horrible' and 14 (7.0%) as 'excruciating'.

The CFA model evaluated using Amos can be best illustrated by a path diagram (Figure 1) of the theoretical model of the JMPQ. According to the model, the sensory dimension (a latent variable) is measured by the first 10 PRI subclasses (observed variables), the affective dimension by the next five subclasses and the evaluative dimension by the 16th subclass. Each of these relationships is illustrated in Figure 1 by an arrow linking the observed variable to the latent variable and is represented statistically by corresponding factor loading, to be explained below. The three dimensions of pain are theoretically intercorrelated, as indicated by the double-headed arrows in Figure 1 and statistically represented by correlation coefficients among the latent variables. The single-headed arrows leading to the observed

variables in the diagram represent the 'unique' (error) portion of each measured variable, ie, that portion not accounted for by the corresponding latent variable. The variance of each of these unique terms is a parameter also estimated by Amos. In the model, the observed 'evaluative' variable does not have a unique or error term because it is the only term designed to measure the evaluative dimension. All of its variance, therefore, is accounted for by the latent variable of the evaluative factor.

As recommended by Byrne et al (23), assessment of fit was based on multiple criteria that reflected statistical, practical and theoretical considerations. Global assessment of fit was based on the likelihood ratio (the χ^2 to degrees of freedom ratio), the Tucker-Lewis Index (TLI) (24), the goodness of fit index (GFI) (25), and the Bentler revised normed comparative fit index (CFI) (26). Interpretations based on the TLI are indicative of the percentage of covariance explained by the hypothesized model, in which values less than 0.90 indicate that the model can be improved substantially (27). The GFI should range between 0 and 1, and high values (greater than 0.9) are associated with a good fit of the model. Adequacy of fit based on CFI values should be greater than 0.90 (26).

CFA

As shown in Table 2, the fit of the hypothesized model was poor from a statistical perspective ($\chi^2[102]=258.16$) and not acceptable from a practical perspective (CFI=0.773, GFI=0.852). In this regard, Byrne et al (23) and Bentler (28) have noted that researchers have been urged not to judge model fit solely on the basis of χ^2 values. Although the coefficient of determination for the first model was excellent (0.996), indicating that the combination of the 16 scales served to measure adequately the factors, the reliability of each observed measure (R^2) with respect to its underlying latent construct ranged from excellent to poor. Specifically, the most reliable scale in the first model was the PRI-3 ($R^2=0.59$), and the PRI-9 was the least reliable ($R^2=0.01$).

SENSITIVITY ANALYSIS

As shown in Figure 1, all factor loadings in the first model (model 1) were statistically significant, with the exception of the PRI-9 subclass scale. Taking into consideration the lack of fit of this model with the observed data, this model was rejected, and the analysis proceeded in an exploratory fashion to specify the sources of misfit in the current model and a series of alternative models. The PRI-9 subclass was retained in post hoc analyses to determine the possible existence of correlated errors with other subclasses and subscales.

To generate alternative models, the constraints for each specified model were relaxed one at a time using the modification indexes as a guide for those parameters in which it made substantive sense to do so. The modification index represents the expected drop in χ^2 if a particular parameter were freely estimated, in which values less than 5.00 show little improvement in fit (23). Previous research with psychological constructs has demonstrated that to generate a well fitting model, it is frequently necessary to allow for correlated

TABLE 1
Data summary for chronic pain sample according to 16 Pain Rating Index (PRI) subclasses of the the McGill Pain Questionnaire, Japanese version

Subclass	Possible range	Number	Range	Chronic pain sample (n=199)		
				Mean	Standard deviation	
PRI-1 (temporal)	0 to 6	167	0 to 6	3.18	1.82	
PRI-2 (spatial)	0 to 3	127	0 to 3	1.42	1.23	
PRI-3 (punctate pressure)	0 to 5	38	0 to 5	0.45	1.11	
PRI-4 (incisive pressure)	0 to 3	14	0 to 3	0.17	0.64	
PRI-5 (constrictive pressure)	0 to 5	80	0 to 5	1.65	2.08	
PRI-6 (traction pressure)	0 to 3	47	0 to 3	0.39	0.82	
PRI-7 (thermal)	0 to 4	22	0 to 4	0.18	0.63	
PRI-8 (brightness)	0 to 4	106	0 to 4	1.98	1.94	
PRI-9 (dullness)	0 to 5	158	0 to 5	3.64	2.10	
PRI-10 (sensory miscellaneous)	0 to 4	57	0 to 4	0.55	1.10	
PRI-11 (tension)	0 to 2	85	0 to 2	0.49	0.61	
PRI-12 (autonomic)	0 to 2	55	0 to 2	0.38	0.67	
PRI-13 (fear)	0 to 3	27	0 to 3	0.39	1.00	
PRI-14 (punishment)	0 to 5	26	0 to 5	0.25	0.84	
PRI-15 (affective miscellaneous)	0 to 2	50	0 to 2	0.30	0.56	
PRI-16 (evaluative)	0 to 5	109	0 to 5	1.94	2.00	

*The number of patients who chose a descriptor in each subclass

TABLE 2
Summary of respecification steps for the McGill Pain Questionnaire, Japanese version, from null to final model

Models	Degrees of freedom	Difference in χ^2	Difference in degrees of freedom		Goodness of fit index	Comparative fit index	Tucker-Lewis index	R^2*
			$\chi^2/\text{degrees of freedom}$	Goodness of fit index				
0 Null	809.35	120		6.75				
1 Initial parameter added	258.16	102	551.19	18	2.53	0.85	0.77	0.73
2 Error correlation between PRI-8 and PRI-1	217.26	101	40.90 [‡]	1	2.15	0.87	0.83	0.80
3 Error correlation between PRI-E [†] and PRI-9	203.58	100	13.68 [‡]	1	2.04	0.88	0.85	0.82
4 Error correlation between PRI-15 and PRI-12	190.43	99	13.15 [‡]	1	1.92	0.89	0.87	0.84
5 Error correlation between PRI-E [†] and PRI-2	177.82	98	12.61 [‡]	1	1.81	0.90	0.88	0.86
6 Error correlation between PRI-7 and PRI-4	168.36	97	9.46 [‡]	1	1.74	0.90	0.90	0.87
7 Error correlation between PRI-9 and PRI-5	158.36	96	10.00 [‡]	1	1.65	0.91	0.91	0.89
8 Error correlation between PRI-12 and PRI-10	150.21	95	8.15 [‡]	1	1.58	0.91	0.92	0.90
9 Error correlation between PRI-5 and PRI-3	143.00	94	7.21 [§]	1	1.52	0.92	0.93	0.92

*The coefficient of determination; [†]The evaluative score of the Pain Rating Index (PRI); [‡] $P<0.001$; [§] $P<0.01$

errors (23,29). This strategy resulted in a final model that allowed error uniqueness between two subclasses or between a subclass and subscale of the same measure to covary. The final model is presented schematically in Figure 2.

Modification indexes

The highest modification indexes (MIs) encountered were: an error covariance between the PRI-8 and the PRI-1 subclasses ($MI=37.25$), the PRI-15 and the PRI-12 subclasses

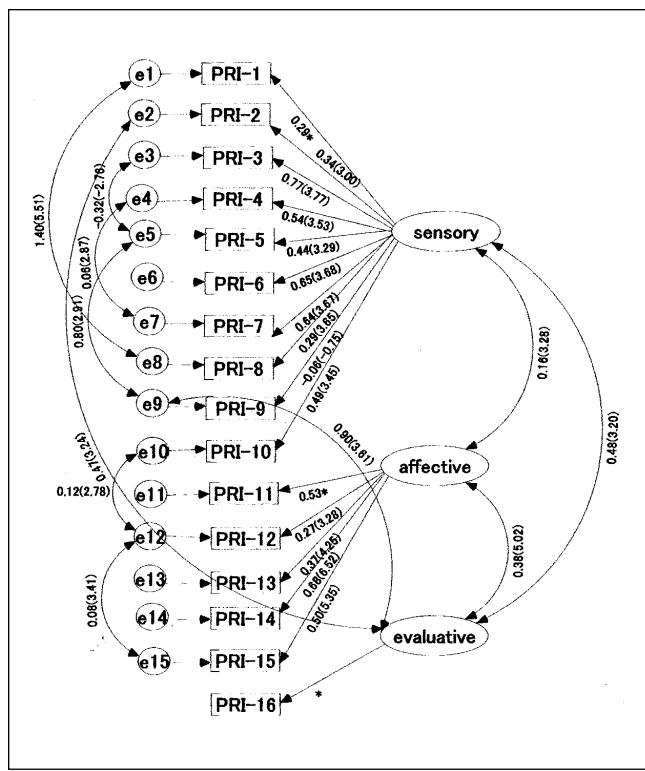


Figure 2) Standardized estimates for three-factor structure of final respecified model underlying the first 16 subclasses of the Japanese version of McGill Pain Questionnaire (MPQ). Values in parentheses are t values of the parameter estimate. Values greater than 1.96 indicate statistical significance ($P < 0.05$). MPQ subscales are sensory, affective and evaluative, and MPQ subclasses are Pain Rating Index (PRI)-1 to PRI-16. Regular boxes represent observed variables and ovals represent latent (unobserved) variables; double-headed arrows represent a pattern of intercorrelation and single-headed arrows leading from the latent constructs to the boxes are regression paths that represent the link from the factors to their respective set of observed variables. *Parameter is fixed to 1.0 for purposes of statistical identification

(MI=12.15), the PRI-7 and the PRI-4 subclasses (MI=10.52), the PRI-9 and the PRI-5 subclasses (MI=9.72), the PRI-12 and the PRI-10 subclasses (MI=8.34), and the PRI-5 and the PRI-3 subclasses (MI=7.56), as well as a covariance between the PRI-E subscale and the PRI-9 subclass (MI=13.05), and the PRI-E subscale and the PRI-2 subclass (MI=10.98).

To assess the extent to which each newly specified model improves over the previous model, the difference in χ^2 between the two models is estimated. Because this differential is itself χ^2 -distributed, with the degrees of freedom being equal to the difference in degrees of freedom between the two models, the significance of this change can be tested statistically. An improvement in model fit is indicated by a large difference in χ^2 value.

Reviewing results related to models 2, 3, 4, 5, 6, 7, 8 and 9 (Table 2), the estimation of each model yielded a significant improvement in fit over its predecessor. Each error covariance, when relaxed, resulted in a statistically significant χ^2 . The final model was acceptable from a practical perspective,

TABLE 3
Intercorrelation between sensory, affective and evaluative subscales of the McGill Pain Questionnaire, Japanese version, with internal consistency reliabilities (Cronbach's alpha) of each subscale and the total scale appearing on the diagonal

	Sensory	Affective	Evaluative	Total score
Sensory	0.63			
Affective	0.56*	0.58		
Evaluative	0.50*	0.47*		
Total score	0.93*	0.74*	0.68*	0.80

* $P < 0.001$

explaining 92% ($TLI=0.92$) of the covariance in the observed data. As noted in Table 2, multiple indicators of fit consistently indicated the adequacy of the final model (eg, $GFI=0.92$, $CFI=0.93$, $TLI=0.92$). As such, model 9 resulted in satisfactory measurement of the pain experience with minimal overlap.

Subscale intercorrelations, discriminant evidence and item correlations

Correlations between subscales and reliability indexes of the total scale and each subscale are shown in Table 3. Subscale intercorrelations were moderate and smaller than reliability indexes.

Regarding the reliability of the JMPQ, internal consistency (Cronbach's alpha) of all subscale scores and the total score on the JMPQ, ranging from 0.58 to 0.80, are listed in Table 3. The total score of the PRI appeared to be satisfactory. The reliability index for the PRI-E cannot be computed because it has only one item.

Significant intercorrelations among the JMPQ scores and other pain rating scale scores (VAS, VRS and NRS), ranging from 0.58 to 0.82, indicate the validity of the JMPQ as a tool for assessing pain ($P < 0.05$).

Intercorrelations between subclasses (the matrix of the item correlations) are also shown in Table 4. Relatively high intercorrelations were demonstrated between subclasses (items).

DISCUSSION

In 1992, Gracely (30) argued in an editorial in *Pain* that exploratory factor analysis methods do not reveal much about the structure of a pain questionnaire when patients are asked to choose items from the scale to describe their pain. Other methods can be used to assess the structure of the questionnaire, but when administered as a symptom checklist, factor analysis tells us more about the characteristics of the patients than the characteristics of the scale. This has been attributed to confusion between semantic and associative meanings. Depending on the studies, between two and seven factors have been reported (4-10). If exploratory factor analysis was used, there is no denying that a three-factor solution may say more about the similarities among the patients of the studies than the 'true' factor structure underlying the MPQ. But CFA was used, which is suited to test a priori postulated factor structure. A well fitting final model was yielded; therefore, theoretically, the results at least supported the hypothesis that

TABLE 4
Intercorrelation between subclasses (matrix of correlations) for the McGill Pain Questionnaire, Japanese version

PRI-1	1.000
PRI-2	0.154* 1.000
PRI-3	0.227** 0.258** 1.000
PRI-4	0.082 0.125 0.448** 1.000
PRI-5	0.162* 0.115 0.189** 0.244** 1.000
PRI-6	0.124 0.302** 0.516** 0.343** 0.207** 1.000
PRI-7	0.207** 0.156* 0.541** 0.499** 0.289** 0.424** 1.000
PRI-8	0.482** 0.245** 0.164* 0.111 0.235** 0.184** 0.123 1.000
PRI-9	0.024 0.011 -0.116 -0.010 0.195** 0.041 -0.055 0.083 1.000
PRI-10	0.133 0.151* 0.308** 0.180* 0.244** 0.411** 0.273** 0.125 -0.038 1.000
PRI-11	0.114 0.249** 0.316** 0.262** 0.310** 0.237** 0.218** 0.181* 0.069 0.169* 1.000
PRI-12	0.057 0.207** 0.121 0.105 0.227** 0.180* 0.123 0.020 0.075 0.315** 0.259** 1.000
PRI-13	0.178* 0.180* 0.239** 0.154* 0.081 0.226** 0.176* 0.275** -0.002 0.201** 0.288** -0.013 1.000
PRI-14	0.134 0.148* 0.518** 0.418** 0.247** 0.456** 0.416** 0.092 -0.083 0.367** 0.337** 0.155* 0.203** 1.000
PRI-15	0.246** 0.159* 0.292** 0.249** 0.317** 0.229** 0.276** 0.246** -0.063 0.303** 0.278** 0.365** 0.220** 0.337** 1.000
PRI-16	0.222** 0.361** 0.322** 0.161* 0.250** 0.335** 0.203** 0.222** 0.208** 0.354** 0.443** 0.240** 0.244** 0.326** 0.252** 1.000
	PRI-1 PRI-2 PRI-3 PRI-4 PRI-5 PRI-6 PRI-7 PRI-8 PRI-9 PRI-10 PRI-11 PRI-12 PRI-13 PRI-14 PRI-15 PRI-16

* $P < 0.05$; ** $P < 0.01$. Sensory values are Pain Rating Index (PRI)-1 to PRI-10, affective values are PRI-11 to PRI-15, and the evaluative value is PRI-16.

the sensory, affective and evaluative subscales of the PRI were representative of the multidimensionality of the pain experience with minimal overlap. Nevertheless, it cannot be denied completely that the a priori three-factor structure merely provides an acceptable fit to the data; other models, such as semantic models, may fit as well or better, as Donaldson (13) pointed out. Additionally, the MPQ is theoretically grounded in the gate control theory of pain (2) and is designed to assess the three dimensions of the pain experience (sensory, affective and evaluative) hypothesized by this theory. Therefore, three subscales of the MPQ were basically assumed to be not completely independent, but mutually inter-correlated and roughly differentiated scales. It is inevitable, therefore, that there are correlated errors among the items (subclasses) of the MPO

Using data based on the responses of the chronic pain patients on the first three subscales of the JMPQ, a well fitting three-factor model was obtained through CFA. This model was established post hoc, by conducting a sensitivity analysis to identify the source of misfit in the initial hypothesized model. Improvement of fit was linked primarily to correlated errors among subclasses for the reasons stated above.

Turk et al (11) concluded that using only the total score of the MPQ was appropriate for pain assessment, because like the present results, the three subscales (sensory, affective and evaluative) were found to be highly intercorrelated. Moreover, reliability indexes were lower than cross-construct correlations. The conclusion of Turk et al (11) was supported by the present results, because contrary to Lowe et al (12), Donaldson (13) and Masedo and Esteve (20), the affective and sensory subscales did not seem to constitute completely independent dimensions of the same construct – pain. As mentioned above and as Holroyd et al (16) pointed out, it cannot be denied that there are relatively high intercor-

relations among three subscales in the Japanese version, similar to the original English version, in view of the process of specifying the final model shown in Table 2, and the result of intercorrelations shown in Tables 3 and 4. Therefore, there is some doubt as to whether three subscales of the JMPQ can actually measure different qualitative aspects of pain at the present time. Indeed, there is skepticism about the discriminant validity of the JMPQ, but one can say that the three subscales of the JMPQ (sensory, affective and evaluative) are theoretically differentiated based on the results of the CFA, but that they are also mutually intercorrelated and not completely independent scales.

To the authors' knowledge, the structure of the JMPQ has not yet been confirmed to be similar to the original format based on back-translation methodology. The present results confirmed the value of the current questionnaire in the field of pain medicine. The authors suggest that a rigorous translation of the MPQ based on back-translation methodology permits the comparison of pain descriptors internationally, despite different languages, thus facilitating international research exchange and communications. The authors expect this study to be the first step toward standardization of the JMPQ and to serve as a cultural bridge in the field of pain medicine between Japan and Canada.

CONCLUSIONS

The results supported the hypothesis that the sensory, affective and evaluative subscales of the PRI are representative of the multidimensionality of the pain experience with minimal overlap. It is suggested that the theoretical structure of the MPQ was kept in the JMPQ used in this study. Therefore, this study may become the first step toward standardization of the Japanese version of the MPQ and serve as a cultural bridge in the field of pain medicine between Japan and English-speaking nations such as Canada.

REFERENCES

1. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975;1:277-99.
2. Melzack R, Wall PD. Pain mechanism: a new theory. *Science* 1965;150:971-9.
3. Prieto EJ, Hopson L, Bradley LA, et al. The language of low back pain: factor structure of the McGill Pain Questionnaire. *Pain* 1980;8:11-9.
4. Kremer EF, Atkinson JH. Pain language: affect. *J Psychosom Res* 1984;28:125-32.
5. Byrne M, Troy A, Bradley LA, et al. Cross-validation of the factor structure of the McGill Pain Questionnaire. *Pain* 1982;13:193-201.
6. Swami DR, Nathawat SS, Solanki RK, Bhojak MM. Factorial validation of McGill Pain Questionnaire. *Ind J Clin Psychol* 1991;18:15-8.
7. Crockett DJ, Prkachin KM, Craig KD. Factors of the language of pain in patients and normal volunteer groups. *Pain* 1977;4:175-82.
8. Burckhardt C. The use of the McGill Pain Questionnaire in assessing arthritis pain. *Pain* 1984;19:305-14.
9. Oostdam EMM, Duivendoorden HJ. Description of pain and the degree to which the complaints fit the organic diagnosis of low back pain. *Pain* 1984;18:171-82.
10. Leavitt F, Garron DC, D'Angelo CM, McNeill TW. Low back pain in patients with and without demonstrable organic disease. *Pain* 1979;6:191-200.
11. Turk DC, Rudy TE, Salovey P. The McGill Pain Questionnaire reconsidered: Confirming the factor structure and examining appropriate uses. *Pain* 1985;21:385-97.
12. Lowe NK, Walker SN, MacCallum RC. Confirming the theoretical structure of the McGill Pain Questionnaire in acute clinical pain. *Pain* 1991;46:53-60.
13. Donaldson GW. The factorial structure and stability of the McGill Pain Questionnaire in patients experiencing oral mucositis following bone marrow transplantation. *Pain* 1995;62:101-9.
14. Brennan AF. The utility of McGill Pain Questionnaire subscales for discriminating psychological disorder in chronic pain patients. *Psychol Health* 1987;1:257-72.
15. McGowan JR, Zebon MA. The relationship between the McGill Pain Questionnaire and functional status in oncology pain patients. *Psychol Health* 1991;5:36-41.
16. Holroyd KA, Holm JE, Keefe FJ, et al. A multicenter evaluation of the McGill Pain Questionnaire: results from more than 1700 chronic pain patients. *Pain* 1992;48:301-11.
17. Lazaro C, Bosch F, Torrubia R, Banos JE. The development of a Spanish Questionnaire for assessing pain: preliminary data concerning reliability and validity. *Eur J Psychol Assess* 1994;10:145-51.
18. Werner O, Campbell D. Translating, working through interpreters, and the problem of decentering. In: Naroll R, Cohen R, eds. *A Handbook of Method in Cultural Anthropology*. New York: Natural History Press, 1970:398-420.
19. Hasegawa M, Hottori S, Ishizaki K, Suzuki S, Goto F. The McGill Pain Questionnaire, Japanese version, reconsidered: Confirming the reliability and validity. *Pain Res Manage* 1996;1:233-7.
20. Masedo AI, Esteve R. Some empirical evidence regarding the validity of the Spanish Version of the McGill Pain Questionnaire (MPQ-SV). *Pain* 2000;85:451-6.
21. Satow A, Nakatani K, Taniguchi S. Analysis of perceptual characteristics of pain describing in words caused by occupational cervicobrachial disorder and similar disease. *Jpn Psychol Res* 1988;30:132-43.
22. Arbuckle JL, Wothke W. *Amos 4.0 User's Guide*. Chicago: SPSS, 1995.
23. Byrne BM, Shavelson R, Muthen B. Testing for the equivalence of factor covariance and mean structure: the issue of partial measurement invariance. *Psychol Bull* 1989;105:456-66.
24. Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika* 1973;38:1-10.
25. Tanaka JS, Huba GJ. A fit index for covariance structure models under arbitrary GLS estimation. *Br J Math Stat Psychol* 1985;38:197-201.
26. Bentler PM. Comparative fit indices in structural model. *Psychol Bull* 1990;107:238-46.
27. Bentler PM, Bonnet DG. Significance tests and goodness-of-fit in the analysis of covariance structure. *Psychol Bull* 1980;88:588-606.
28. Bentler PM. Multivariate analysis with latent variables: Causal modeling. *Annu Rev Psychol* 1980;31:419-56.
29. Tanaka JS, Huba GJ. Confirmatory hierarchical factor analyses of psychological distress measures. *J Pers Soc Psychol* 1984;46:621-35.
30. Gracey RH. Evaluation of multi-dimentional pain scales. *Pain* 1992;48:297-300.



The Hindawi logo consists of two interlocking circles, one blue and one green, forming a stylized infinity or double helix symbol.

Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

