PAIN 01667

Assessment of the components of observed chronic pain behavior: the Checklist for Interpersonal Pain Behavior (CHIP)

Johan W.S. Vlaeyen ^{a,b}, Dia F.M. Pernot ^a, Ank M.J. Kole-Snijders ^{a,c}, Joop A. Schuerman ^{a,b,c},
Hugo Van Eek ^c and Nico H. Groenman ^b

^a Institute for Rehabilitation Research, Hoensbroek,

^b Department of Medical Psychology, University of Limburg, Maastricht,

and ^c Lucas Foundation for Rehabilitation, Hoensbroek (The Netherlands)

(Received 28 November 1989, revision received 3 January 1990, accepted 24 May 1990)

Summary This article describes the development of the Checklist for Interpersonal Pain Behavior (CHIP), an observation scale which assesses overt pain behavior. The study is an extension of an earlier study [29] in which the dimensions and components of observed chronic pain behavior were examined. A broad definition of pain behavior is chosen (interpersonal pain behavior), namely the interaction between the pain patient and his/her direct environment. The list of pain behaviors, taken from the earlier study, has been transformed into a 78-item global rating scale to be used by nurses to quantify observed pain behavior in a clinical setting. Six studies examine the factor structure and the psychometric properties of this behavioral observation method. In the first study, 6 internally reliable factors are derived using factor analytic techniques from a sample of 152 chronic pain patients. They are labeled as: 'distorted mobility,' 'verbal complaints,' 'non-verbal complaints,' 'nervousness,' 'depression' and 'day sleeping.' Internal consistency of all factors, except 'day sleeping' was excellent. The following studies show that the CHIP is sufficiently reliable and valid. After a discussion on the advantages of this observation scale, the conclusion seems justified that the CHIP is a useful tool in pain assessment that can easily be used by nurses.

Key words: Chronic pain; Pain behavior; Behavioral assessment; Observation

Introduction

In the literature on chronic pain conceptualization, assessment and treatment, pain is more frequently approached as a multifaceted condition involving three relatively independent response systems: gross motor, cognitive and physiological [1,4,14,19,20,24,29,30]. The first response system refers to motoric communications of suffering through overt pain behaviors. In relation to the

second response system, pain patients simultaneously assign personally relevant meanings or causal attributions [9,31] to their pain problem, and they evaluate options for coping with pain (self-efficacy expectations) [6]. The third response system refers to psychophysiological correlates of pain, including changes in cerebral evoked potentials, muscle tension and skin resistance [8].

Under the influence of Fordyce's theory of operant conditioning of pain behaviors [10], a great deal of clinical research has dealt with the modification of overt expressions of pain and suffering (summarized by Linton [17]). The overall result is that pain behaviors can be extinguished successfully and that healthy behaviors can be

Correspondence to: Johan W.S. Vlaeyen, Institute for Rehabilitation Research, Zandbergsweg 111, 6432 CC Hoensbroek, The Netherlands.

shaped through operant conditioning programs. However, in order to evaluate the effectiveness of these programs, reliable and valid instruments are needed to measure pain behaviors. One recurrent problem that characterizes attempts to assess pain behaviors involves the characteristics of the 'pain behavior' construct. There is little agreement about the range of behaviors that should be assessed [26]. This problem has stimulated researchers to examine empirically the dimensions and components of the pain behavior construct [22,26,29]. These studies provide an empirically derived and methodologically sound basis for the construction of a pain behavior measure.

In this context, Vlaeyen et al. [29] replicated the study of Turk et al. [26] and found that pain behavior — in the broad sense of the term — can be characterized along three dimensions: 'withdrawal-approach,' 'high arousal-low arousal,' and 'visible-audible.' In addition, 9 components were identified: (1) anxiety; (2) attention seeking; (3) verbal pain complaints; (4) medication use; (5) general verbal complaints; (6) distorted posture and mobility; (7) fatigue; (8) insomnia; (9) depressive mood.

In this study [29] the original list of pain behaviors was generated by nurses who carefully observed chronic back pain patients. The 78 items found were subsequently sorted by 96 psychologists. It was suggested that the same items could provide a useful basis for the development of an observational measurement technique for chronic pain.

The present study is designed to be an extension of the above-mentioned study, and is an attempt to transform the original pool of 78 items into a reliable and valid observation scale for pain behavior. A device which directly measures pain behavior in the broad sense, as identified in the above-mentioned study, would represent a useful addition to the diagnostic protocol of pain management centers. For the purpose of clarity, and in order to differentiate between the narrow and broad definition of 'pain behavior,' the term 'interpersonal pain behavior' is introduced, referring to the broad definition.

Six studies will be presented. The first study examines the factor structure of this new observa-

tional instrument. Subsequently, 2 studies are described which focus on the reliability of the CHIP. Finally, 3 studies examine the validity of the instrument. All of the studies were carried out with patients admitted to an inpatient rehabilitation program at the Department of Rheumatology of the Lucas Foundation for Rehabilitation in Hoensbroek, The Netherlands.

Study I

The 78 items derived from the previous study [29] were transformed into a checklist in which they are followed by a 5-point Likert scale. By choosing one of the 5 alternatives ('never observed,' 'seldom observed,' 'occasionally observed,' 'frequently observed' and 'almost always observed') the frequency of occurrence of each item could be rated. This new observational method was named 'Checklist for Interpersonal Pain behavior (CHIP).' The purpose of Study I was to analyze the factor structure of the CHIP and to evaluate the characteristics of the factors found.

Method

Subjects

Eighty-two female and 70 male chronic pain patients consecutively admitted to an inpatient rehabilitation program were carefully observed during their first two weeks of admission by one of 12 nurses. The patients' mean ages were 47.4 (S.D. = 14.9) and 39.6 (S.D. = 13.6) years, respectively. The sample was diagnostically heterogeneous. Seventy-two patients (47.4%) were diagnosed as having chronic benign back pain (CBBP), 58 (38.2%) suffered from rheumatoid arthritis, 12 (7.8%) suffered from spondylitis ankylopoetica, and 10 (6.6%) were diagnosed as having fibromyalgia. One hundred and twenty-four patients (78%) were admitted for chronic back pain problems.

Procedure

The CHIP was used by one of a group of 12 nurses to quantify observed pain behaviors. The observation took place during the first week of

admission, after which the checklist was filled in for each patient. The 152 checklists were subsequently subjected to frequency counts and the frequency distribution was analyzed. Only those items with a frequency distribution approximating the normal distribution were accepted. Items with a skewness < -1 and > 1 , and items which were not observed in more than 10% of the cases, were rejected. A principal component factor analysis with Varimax rotation was applied to the remaining items. Items with a communality < 0.30 and items with a factor loading < 0.40 on all of the

factors were also rejected. Finally, an analysis of variance (ANOVA) was carried out in order to examine the relationship between factor scores and individual differences including sex, age (under vs. over the median = 43.5 years), and diagnosis (CBBP vs. other diagnoses group).

Results and discussion

Analysis of the frequency distribution of the 78 items revealed that only 34 items were quasi-normally distributed. Six items were not observed in

TABLE I

VARIMAX ROTATED FACTOR MATRIX AND RELIABILITY COEFFICIENTS FOR THE INDIVIDUAL ITEMS (TRANSLATED FROM THE DUTCH)

Numbers refer to the original behavior list. A, intra-rater reliability; B, inter-rater reliability (Pearson product-moment correlation coefficients for both A en B).

Item	CHIP factor						A	B
	1	2	3	4	5	6		
33. Walks with abnormal gait	0.83						0.60 **	0.85 **
43. Stands in unusual posture	0.80						0.57 **	0.53 **
17. Moves rigidly and stiffly, with stretched limbs	0.75						0.62 **	0.50 **
64. Walks guardedly and carefully	0.61						0.70 **	0.65 **
56. Has bad mobility	0.60						0.49 **	0.51 **
63. Drags feet when walking	0.44						0.30	0.43 *
71. Complains about pain		0.70					0.57 **	0.73 **
26. Gives lively descriptions of the pain		0.63					0.55 **	0.63 **
76. Says that the pain has not diminished		0.59					0.72 **	0.74 **
5. Mentions having more pain after therapy		0.55					0.67 **	0.22
13. Says that she/he is tired		0.47					0.48 **	0.48 **
21. Changes from one buttock to the other while sitting			0.79				0.12	0.43 *
78. Changes position frequently while sitting			0.77				0.42 *	0.58 **
18. Rubs painful parts of the body			0.56				0.47 **	0.60 **
24. Groans, moans or sighs			0.51				0.40 *	0.59 **
15. Restless and nervous				0.75			0.48 **	0.80 **
1. Tense				0.70			0.57 **	0.57 **
54. Worries				0.54			0.36 *	-0.08
4. Is quiet and withdrawn					0.84		0.51 **	0.60 **
29. Isolates her/himself					0.73		0.72 **	0.63 **
11. Appears drowsy and sleepy					0.42		0.49 **	0.50 **
67. Appears blue, down				0.49	0.41		0.33	0.38 *
2. Falls asleep during the day						0.89		
14. Takes a nap after lunch						0.61		

* $P < 0.01$, ** $P < 0.001$.

more than 10% of the cases. A principal component factor analysis was carried out on the remaining 28 items. Using the criterion of factors with eigenvalues > 1 to determine the number of factors to be extracted, 6 factors were obtained which cumulatively accounted for 67.1% of the total variance. To maximize the variance of the squared factor loadings, the original 6-factor solution was subsequently subjected to an orthogonal Varimax rotation. Item loadings ≥ 0.40 were considered significant. Using this criterion, item 67 showed high loadings on both factor 4 and factor 5. Therefore, Pearson correlation coefficients between item 67 and factors 4 and 5 were calculated (0.72 and 0.81, respectively, both $P < 0.001$). As the correlation with factor 5 was greater, item 67 was assigned to this factor. Factors were labeled following the common procedure of giving the greatest consideration to items with high loadings on each factor. The following labels were assigned to the factors: 'distorted mobility,' 'verbal complaints,' 'non-verbal complaints,' 'nervousness,' 'depression,' and 'day sleeping.' Table I shows the items with their respective factor loadings. Taking into account the reduced item pool — only 36% of the original items could be included in the factor analysis — the factor structure does correspond fairly well with the clusters found in the earlier study.

Characteristics of the CHIP factors including factor label, number of items per factor, score range, skewness, factor means and standard deviations, and Cronbach's alpha are presented in Table II. All CHIP factors appeared to be quasi normally distributed (skewness < 1). For factors 1

to 5, Cronbach's alpha, the widely accepted measure for internal consistency of a test, ranged from 0.83 to 0.89, suggesting excellent internal consistency. Because of its poor internal consistency and its limited number of items, CHIP factor 6 was excluded from further analysis.

The ANOVA testing the relationship between factor scores and individual differences showed significant differences between male and female patients for all factors ($P < 0.01$). This finding suggests that female pain patients showed more interpersonal pain behavior than the male pain patients. For factors 3 and 5 significant differences emerged between the two age groups ($F = 4.86$, $P = 0.03$ and $F = 4.15$, $P = 0.04$, respectively), suggesting that older pain patients displayed more non-verbal complaints and showed more depressed behavior than younger pain patients. For factors 2 and 3, ANOVA revealed significant differences between the diagnosis groups ($F = 4.01$, $P = 0.048$ and $F = 5.84$, $P = 0.02$, respectively), suggesting that CBBP patients displayed more verbal and non-verbal pain behavior than the other diagnoses group. Interestingly, for factors 1, 4, and 5, significant 2-way interactions emerged between sex and diagnosis. A post-hoc t test revealed that within the female group the CBBP patients scored significantly higher than the other diagnoses group on these factors ($t = 2.54$, $P < 0.01$; $t = 1.75$, $P < 0.05$; and $t = 3.72$, $P < 0.001$, respectively). This was not the case for the male patient group. For factor 1, male CBBP patients even scored significantly lower than the male other diagnoses group ($t = -2.23$, $P < 0.025$).

TABLE II
CHARACTERISTICS OF THE CHIP FACTORS

Factor label	Number of items	Range	Skewness	Factor mean	Factor S.D.	Cronbach's alpha	Intra-rater reliability	Inter-rater reliability
1. Distorted mobility	6	0–24	0.347	9.1	7.1	0.89	0.74 **	0.83 **
2. Verbal complaints	5	0–20	0.666	7.1	4.5	0.83	0.70 **	0.77 **
3. Non-verbal complaints	4	0–16	0.725	5.4	3.8	0.85	0.49 **	0.67 **
4. Nervousness	3	0–12	0.204	4.8	3.1	0.85	0.64 **	0.69 **
5. Depression	4	0–16	0.672	4.6	3.6	0.83	0.45 **	0.65 **
6. Day sleeping	2	0– 8	–0.052	3.2	2.3	0.74		

* $P < 0.01$, ** $P < 0.001$.

These results suggest that, in general, female pain patients display more interpersonal pain behavior than male pain patients, and that this is more pronounced for female CBBP patients, who appear to display more distorted mobility, nervousness and depressive behaviors than female patients with other diagnoses. Moreover, CBBP patients show more verbal and non-verbal pain behavior than patients with other pain diagnoses. This means that if norms are to be developed, different norm groups including both sexes, age groups and diagnosis groups will have to be taken into consideration.

In order to utilize the 22-item CHIP as an observational measurement instrument, the following studies examined both reliability and validity of the CHIP factors. The patients serving as subjects of these studies were classified by their physicians as either having minimal organic findings or displaying pain complaints disproportionate to the demonstrable organic basis of their pain.

Study II

Reliability analysis of the CHIP means assessing internal consistency, intra-rater reliability, and inter-rater reliability for all internally consistent CHIP factors. Study II assessed the intra-rater reliability of the individual CHIP items and CHIP factors by examining correlations between 2 successive observations.

Method

Subjects

Thirteen male and 33 female chronic low back pain patients, admitted for an inpatient rehabilitation program, were observed at two different points in time. Their mean ages were 40 years (range 28–52 years) and 44 years (range 24–60 years) respectively. The mean duration of their pain complaints was reported to be 11.9 years, ranging between 1 and 37 years.

Procedure

The observation took place during 2 pain screening days being held before admission to the

rehabilitation centre. The time interval between the 2 observations was 2 weeks. During the screening days, a trained research assistant scored the CHIP while exercise tolerance levels were assessed by means of a standardized instruction which was based on the pain-rest contingency principle: 'Carry on until pain or physical discomfort makes you stop the activity.' Both dynamic activities (walking, riding on a bicycle ergometer, arm and leg rowing, and shifting objects from one place to another) and static activities (sitting and standing) were requested from the patients. Pearson product-moment correlation coefficients were calculated between the two successive administrations for the individual CHIP items and for the 5 CHIP factor scores.

As the subjects in this study also participated in study V and study VI, the level of significance was conservatively set at $P = 0.01$.

Results and discussion

As can be seen in Table I (column A), only three items (21, 63 and 67) appear to be insufficiently reliable. Table II displays the reliability coefficients for the 5 CHIP factors, which are all significant. These results suggest that the CHIP is stable across time, at least when being used by trained observers who are able to carefully observe pain patients who are engaged in physical activities.

Study III

Study III evaluated the inter-rater reliability of the CHIP, by examining the correlations between scores of the individual CHIP items and CHIP factors obtained by two independent observers.

Method

Subjects

The sample of this study consisted of 37 female and 10 male chronic low back pain patients with mean ages 44.8 years (range 24–64 years) and 35.3 years (range 27–52 years) respectively. The mean duration of their pain complaints was reported to

be 12.2 years ranging between 6 months and 31 years.

Procedure

The procedure was similar to the one in Study II. The inter-rater reliability analysis was carried out with 2 trained research assistants who independently observed the patients during a measurement day on which exercise tolerance levels were assessed as described in the previous study. The observation took place 2 weeks after admission.

Results and discussion

Pearson product-moment correlation coefficients between the scores for the individual CHIP items revealed significant inter-rater reliability for all items except items 5 and 54 (Table I, column B). Correlations for items 63, 21 and 67 are significant but low. For the 5 CHIP factor scores, significant correlations are found (Table II). These results suggest that the inter-rater reliability of the CHIP is acceptable. However, because of their weak intra- and inter-rater reliability, removal of items 21, 54, 63, and 67 from the CHIP might be considered in future research (Removal of items 21, 54, 63 and 67 affects internal consistency of CHIP factors 1, 3, 4 and 5. Cronbach's alpha changes from 0.89, 0.85, 0.85, and 0.83 to 0.88, 0.77, 0.85, and 0.79 respectively).

Study IV

The heterogeneous character of the CHIP factors makes it difficult to examine validity in a simple way. In fact, validity for each individual factor should be assessed. Nevertheless, one form of validity, content validity, is assumed to be sufficient for all CHIP factors. Content validity is determined by the adequacy with which an instrument samples the behavioral domain of interest [5]. As the CHIP is directly derived from an empirical study that examined the dimensions and components of pain behavior, it may be concluded that content validity has already been proven.

Study IV assessed the criterion-related validity of CHIP factors 1 and 3. Criterion-related valid-

ity, also referred to as concurrent validity, refers to the usefulness or accuracy of scores in predicting some performance criterion or in substituting for some other, more costly, measurement procedure [2,21]. In study IV, the extent to which the CHIP scores correspond with the number of pain behaviors counted on videotapes is examined. Correlations between the CHIP scores and another existing observational instrument are also analyzed.

Method

Subjects

In this study, the subject sample consisted of 15 female and 4 male chronic low back pain patients with mean ages 43.7 (range 24–55 years) and 38.7 years (range 29–52 years), respectively. Thirty video recordings of these patients were made while their functional skills were being assessed using the pain-rest contingency principle as described in Study II.

Procedure

The Audiovisual Taxonomy for Assessing Pain Behavior (ATAPB), developed by Keefe et al. [15], and the UAB Pain Behavior Scale (PBS), developed by Richards et al. [23] were chosen as concurrent measures for the CHIP factors 'distorted mobility' and 'non-verbal pain complaints'.

The ATAPB allows 5 behavior categories to be observed, namely sighing, bracing, guarding, rubbing, grimacing, and a total score. Psychometric analysis of the ATAPB [16] showed that the total score has an acceptable reliability, and that the individual pain behaviors do not occur at sufficient frequencies to be considered alone. In this study, it was decided to compute both the individual behavior scores and the total scores.

The PBS consists of 10 target pain behaviors assumed by Richards et al. [23] to be reliably observed in the chronic pain population. The PBS is also validated for outpatient use [7].

Both CHIP and PBS ratings were obtained on the same days that video recording took place. The audiovisual taping was done individually and patients were aware of the videotaping. According to Keefe et al.'s study, chronic pain patients were

videotaped while involved in physical activities as described in Study III. Each videotaped session was divided into segments of 30 sec. To ensure consistency in the intervals, a recorded audio signal was dubbed onto each videotape. Depending on the length of the video recording, a different number of segments was obtained. After each segment, 10 sec were used to rate the ATAPB. Scores were derived for each category by adding the total number of occurrences of each coding category over the observation session. In order to obtain comparable values, this score was divided by the number of segments and then multiplied by 100. A total pain behavior score consisted of the sum of each category score. For the total pain behavior score interobserver agreement between two trained observers who both scored 15 video recordings was 0.94 ($P < 0.001$). One of them completed the 30 recordings, and the data derived from her observations were included in further analyses. Because of the relatively small sample size, the Kendall correlation coefficient Tau is chosen in this study.

Results and discussion

Correlations between scores for CHIP factors 'distorted mobility' and 'non-verbal complaints,' on the one hand, and ATAPB scores and PBS scores on the other, were calculated; these are summarized in Table III. It should be noted that similar to the study of Kleinke and Spangler [16], 'sighing' and 'rubbing' occurred at very low frequencies. The two CHIP factors 'distorted mobility' and 'non-verbal complaints' correlate significantly with both the ATAPB total scores and the PBS scores, indicating a good criterion-related

validity. Of interest is that the PBS displays a pattern of significant correlations with the ATAPB scores similar to the CHIP factors 1 and 3.

Study V

This study assessed the construct validity of CHIP factors 1, 2, and 3 by examining the relationship between the CHIP scores and measures of health behavior. Pain behaviors are conceptualized to be incompatible with health behaviors such as physical performance [11]. Therefore, it was hypothesized that negative correlations are to be found among CHIP factor-scores and measures of physical performance.

Method

Subjects

In Study V, the subjects were the same as those involved in Study II.

Procedure

The procedure is similar to the one described in Study II, but on the second observation day, exercise tolerance levels were recorded. The patients were asked to walk, to ride a stationary bicycle, and to perform desk work in both standing and sitting positions. To assess exercise tolerance, patients were asked to perform the target activity '... until pain or physical discomfort causes you to wish to stop.' The distance walked, and the time spent riding the bicycle, standing and sitting were recorded. Pearson product-moment correlations were calculated among these measures of health behavior and scores on CHIP factors 1, 2, and 3.

TABLE III
CRITERION-RELATED VALIDITY FOR CHIP FACTORS 1 AND 3

Kendall correlation coefficient Tau, with ATAPB, Audiovisual Taxonomy for Assessing Pain Behavior; PBS, Pain Behavior Scale.

CHIP factor	ATAPB category						PBS
	Sighing	Bracing	Guarding	Rubbing	Grimacing	Total	
CHIP 1	0.09	0.65 ***	0.50 ***	0.21	0.11	0.43 **	0.52 ***
CHIP 3	0.27	0.79 ***	0.75 ***	0.25	0.53 **	0.81 ***	0.81 ***
PBS	0.22	0.58 ***	0.56 ***	0.28 *	0.41 **	0.62 ***	—

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Results and discussion

As displayed in Table IV, an interesting differential pattern of significant correlations emerged among CHIP factor scores and exercise tolerance measures. All of the correlations are significant, except for CHIP factor 2 (verbal complaints). Correlations with walking and bicycle riding are non-existent. This finding suggests that verbal complaints may also occur with pain patients who walk long distances and spend a longer time on the bicycle. In accordance with the hypothesis, and with the item content of CHIP factor 1, correlations with this factor are more pronounced for the exercises involving mobility (walking and bicycle riding) as compared to the other, more stationary exercises (-0.51 , -0.60 vs. -0.37 , -0.31). A similar pattern is seen for CHIP factor 3. Correlations are greater for activities involving sitting (bicycle riding and sitting), as compared to the exercises involving walking and standing (-0.38 , -0.62 vs. -0.19 , -0.27 , respectively). This finding is also in accordance with the item content of CHIP factor 3.

The findings clearly support the construct validity of CHIP factors 1 and 3. For factor 2, further validity assessment will be necessary.

Study VI

Study VI examined the relationship between CHIP factors 4 and 5 and relevant self-report

TABLE IV
CRITERION-RELATED VALIDITY FOR CHIP FACTORS 1, 2 AND 3

Pearson product-moment correlation coefficient with measures of health behaviors 'walking,' 'bicycle riding,' 'standing' and 'sitting.'

CHIP factor	Walking	Bicycle	Standing	Sitting
CHIP 1	-0.51^{**}	-0.60^{**}	-0.37^*	-0.31
CHIP 2	0.07	-0.11	-0.30	-0.38^*
CHIP 3	-0.19	-0.38^*	-0.27	-0.62^{**}

* $P < 0.01$, ** $P < 0.001$.

TABLE V

CRITERION-RELATED VALIDITY FOR CHIP FACTORS 4 AND 5

Pearson product-moment correlation coefficients (r_{xy}) and partial correlations ($r_{xy.z}$) among CHIP factors and self-report measures MMPI-depression and WAS.

	MMPI-D		WAS	
	r_{xy}	$r_{xy.z}$	r_{xy}	$r_{xy.z}$
CHIP 1	0.23	0.22	0.07	-0.01
CHIP 2	0.23	0.09	0.37 *	0.31
CHIP 3	0.24	0.13	0.30	0.25
CHIP 4	0.45 **	0.32	0.49 **	0.41 *
CHIP 5	0.47 **	0.42 *	0.27	0.07
WAS	0.40 *			

* $P < 0.01$, ** $P < 0.001$.

measures. Although there are many studies in the pain assessment literature in which intercorrelations are investigated between self-ratings and observations, such correlations do not 'validate' any of the measures. Nevertheless, if significant correlations emerge, they might support the labeling of the CHIP factors.

Method

Subjects

In this study, the subjects were the same as those involved in Study II.

Procedure

As no reliable Dutch observational instruments were available as concurrent measures for CHIP factors 4 and 5, it was decided to choose relevant self-report measures. The following measures were selected: (a) The MMPI-Depression Scale (MMPI-D) [13] is considered a valid self-rating of depression. (b) The Welsh Anxiety Scale (WAS) [32], derived from the MMPI [31] was considered as a general measure of nervousness.

The observation occurred as described in Study II. For this study, however, only the data of the second observation day are used. At this moment the MMPI was also available, of which the MMPI-D scale and the WAS raw scores were calculated. Correlations were calculated among

these self-report measures and the CHIP factors 1–5.

Results and discussion

The differential pattern of significant correlations is shown in Table V. Significance emerged only for CHIP factors 4 and 5, suggesting that these factors indeed capture the observational component of depression and nervousness. The significant correlation between CHIP factor 4 and MMPI-D might also be due to the existing correlation between MMPI-D and WAS. Therefore partial correlations were calculated, controlling for WAS and MMPI-D respectively. Significance of the correlations emerged only in the hypothesized direction, thereby confirming the validity of CHIP factors 4 and 5.

These results, however, should be regarded with caution. As a different kind of assessment is used (observation vs. self-report), lower correlations might be expected than in the case where only self-report or only observational measures are correlated [12].

General discussion

Based on the results of an earlier study [29] in which the components of observed chronic pain behavior were identified, the present study offered an operationalization of the broad ‘interpersonal pain behavior’ construct in the form of the CHIP. This paper also examined the psychometric properties of this new observational instrument. Both reliability and validity of the main factors appear promising.

Unlike self-reports of the pain experience, which rely on the patient him/herself to provide data, the CHIP represents a global rating scale that provides data resulting from direct observation. The observation occurs in a therapeutic environment and can be made by nurses who frequently interact with the patient outside structured therapy sessions. The advantages of the CHIP are that it is based on empirically derived items and that it is relatively short and easy to use. Because of its simplicity, the CHIP might be suitable for use by

the patient’s spouse in the home environment as well.

Compared to existing observation methods, the CHIP might provide a more comprehensive picture of the pain patient. A recent meta-analysis of non-medical treatments for chronic pain [18] indicated that psychological treatments reliably affect mood and subjective symptom ratings. The CHIP offers 2 observational measures related to mood (nervousness and depression).

One disadvantage is that a certain degree of inference is required to complete a rating. The rater must observe the patient and, through inference, determine which rating is most appropriate. Moreover, the CHIP, like most global rating scales, may be susceptible to observer bias. As a consequence, explicit training, during which a consensus must be reached with regard to the specific meaning of the individual items, is a necessary prerequisite for its application. Finally, the CHIP has not yet been examined with regard to its sensitivity as a measure of patient progress during a pain management program.

A weakness of this study concerns the results of study V, which do not fully support the validity of the factors ‘nervousness’ and ‘depression’. A similar study, but based solely on observational concurrent measures (when they become available) is recommended. Furthermore, CHIP factor ‘verbal complaints’ awaits further validation.

In recent publications [25,27,29] researchers have stated that it is necessary to establish the presence or absence of respondent, operant, and cognitive contributions to the pain problem, and that treatments should be based on these ‘patient profiles.’ In spite of its disadvantages, the CHIP may be a useful instrument in identifying chronic pain patients whose pain problem can be characterized as being mainly overt-motoric, as compared to more cognitive or physiological pain problems. During a pre-treatment baseline observation period, high scores on the CHIP may indicate that the patient is a good candidate for an operant pain management program aimed at reducing pain behaviors and restoring healthy behaviors. Furthermore, many disadvantages can be offset and the advantages retained if multiple measures are employed. Observational data should

be integrated with other data into a Multiaxial Assessment of Pain [28] that should enhance our understanding of chronic pain.

Acknowledgements

The authors wish to express their gratitude to A.H. Bakker-Boerrigter, MD and the staff of the Department of Rheumatology of the Lucas Foundation for Rehabilitation in Hoensbroek, The Netherlands. Thanks are also extended to J. van Houtem for his statistical advice, and to T. Fasotti-Dumont, C. Magnée and B. Noorloos for their research assistance.

References

- 1 Bakal, D.A., Demjen, S. and Koganov, J., The continuous nature of headache susceptibility, *Soc. Sci. Med.*, 12 (1984) 1305–1311.
- 2 Barrios, B. and Hartmann, D.P., The contributions of traditional assessment: concepts, issues and methodologies. In: R.O. Nelson and S.C. Hayes (Eds.), *Conceptual Foundations of Behavioral Assessment*, The Guilford Press, New York, 1986.
- 3 Beck, A.T., *Depression: Causes and Treatment*. University of Philadelphia Press, Philadelphia, 1967.
- 4 Cautela, J.R., The use of covert conditioning in modifying pain behavior. *J. Behav. Ther. Exp. Psychiatry*, 8 (1977) 45–52.
- 5 Cronbach, L.J., Test validation. In: R.L. Thorndike (Ed.), *Educational Measurement* (2nd edn.), American Council on Education, Washington, DC, 1971.
- 6 Dolce, J.J., Self-efficacy and disability beliefs in behavioral treatment of pain, *Behav. Res. Ther.*, 4 (1987) 289–299.
- 7 Feuerstein, M., Greenwald, M., Gamache, M.P., Papciak, A.S. and Cook, E.W., The Pain Behavior Scale: modification and validation for outpatient use, *J. Psychopathol. Behav. Assess.*, 7 (1985) 301–315.
- 8 Flor, H., Turk, D.C. and Birbaumer, N., Assessment of stress-related psychophysiological reactions in chronic back pain patients, *J. Consult. Clin. Psychol.*, 53 (1985) 354–364.
- 9 Flor, H. and Turk, D.C., Chronic back pain and rheumatoid arthritis: predicting pain and disability from cognitive variables, *J. Behav. Med.*, 11 (1988) 251–265.
- 10 Fordyce, W.E., *Behavioral Methods in Chronic Pain and Illness*, Mosby, St. Louis, MO, 1976.
- 11 Fordyce, W.E., Shelton, J.L. and Dundore, D.E., The modification of avoidance learning in pain behaviors, *J. Behav. Med.*, 5 (1982) 405–414.
- 12 Gannon, L.R. and Haynes, S.N., Cognitive-physiological discordance as an etiological factor in psychophysiological disorders. In: S. Rachman (Ed.), *Adv. Behav. Res. Ther.* Vol. 8, Raven Press, New York, 1986, pp. 223–236.
- 13 Hathaway, S.R. and McKinley, J.C., *Minnesota Multiphasic Personality Inventory* (revised edn.), Psychological Corporation, New York, 1967.
- 14 Keefe, F.J., Behavioral assessment and treatment of chronic pain, current status and future directions, *J. Consult. Clin. Psychol.*, 6 (1982) 896–911.
- 15 Keefe, F.J., Crisson, J.E. and Trainor, M.T., Observational methods for assessing pain: a practical guide. In: J.A. Blumenthal and D. McKee (Eds.), *Applications in Behavioral Medicine and Health Psychology: A Clinician's Sourcebook*, Professional Resource Exchange, Sarasota, FL, 1987, pp. 67–94.
- 16 Kleinke, C.L. and Spangler, Jr., A.S., Psychometric analysis of the audiovisual taxonomy for assessing pain behavior in chronic back pain patients, *J. Behav. Med.*, 11 (1988) 83–94.
- 17 Linton S.J., Behavioral remediation of chronic pain: a status report, *Pain*, 24 (1986) 125–141.
- 18 Malone, M.D. and Strube, M.J., Meta-analysis of non-medical treatments for chronic pain, *Pain*, 34 (1988) 321–344.
- 19 Mayer, T.G., Gatchel, R.J., Kishino, N., Keeley, J., Mayer, H., Capra, P. and Mooney, V., A prospective short-term study of chronic low back pain patients utilizing novel objective functional measurement, *Pain*, 25 (1986) 53–68.
- 20 Melamed, B.C., Health intervention: collaboration for health and science. In: B.L. Hammonds and C.J. Schreier (Eds.), *Psychology and Health*, American Psychological Association, Washington, DC, 1984.
- 21 Mischel, W., *Personality and Assessment*, Wiley, New York, 1986.
- 22 Philips, H.C. and Jahanshahi, M., The components of pain behaviour report, *Behav. Res. Ther.*, 2 (1986) 117–125.
- 23 Richards, J.S., Nepomuceno, C., Riles, M. and Suer, Z., Assessing pain behavior: the UAB pain behavior scale, *Pain*, 12 (1982) 393–398.
- 24 Sanders, S.H., Behavior assessment and treatment of chronic pain: appraisal and current status. In: M. Hersen, R.M. Eisler and P.M. Miller (Eds.), *Progress in Behavior Modification*, Vol. 8, Academic Press, New York, 1979, pp. 249–291.
- 25 Syrjala, K.L. and Chapman, C.R., Measurement of clinical pain: a review and integration of research findings. In: C. Benedetti et al. (Eds.), *Advances in Pain Research and Therapy*, Vol. 7, Raven Press, New York, 1984.
- 26 Turk, D.C., Wack, J.T. and Kerns, R.D., An empirical examination of the 'pain-behavior' construct, *J. Behav. Med.*, 8 (1985) 119–130.
- 27 Turk, D.C. and Flor, H., Pain > pain behaviors: the utility and limitations of the pain behavior construct, *Pain*, 31 (1987) 277–295.
- 28 Turk, D.C., Rudy, T.E. and Stieg, R.L., The disability determination dilemma: toward a multiaxial solution, *Pain*, 34 (1988) 217–229.

- 29 Vlaeyen, J.W.S., Van Eek, H., Schuerman, J.A. and Groenman, N.H., Dimensions and components of observed chronic pain behavior, *Pain*, 31 (1987) 65–75.
- 30 Vlaeyen, J.W.S., Snijders, A.M.J., Schuerman, J.A., Van Eek, H., Groenman, N.H. and Bremer, J.J.C.B., Chronic pain and the three-systems model of emotions. A critical examination, *CRC Crit. Rev. Phys. Rehab. Med.*, 2 (1989) 67–76.
- 31 Weiner, B. and Graham, S., Attributional approach to emotional development. In: C.E. Izard, J. Kagan and R.B. Zajonc (Eds.), *Emotions, Cognition and Behavior*, Cambridge University Press, London, 1984.
- 32 Welsh, G.S., MMPI profiles and factor scales A and R, *J. Clin. Psychol.*, 21 (1965) 43–47.