A Revised Taxonomy of Patients with Chronic Pain

Christine E. Sheffer, PhD,* Julie A. Deisinger, PhD,† Jeffrey E. Cassisi, PhD,‡ and Kenneth Lofland, PhD§

*Department of Health Behavior and Health Education, College of Public Health, University of Arkansas for Medical Sciences, Little Rock, Arkansas; †Department of Psychology, Saint Xavier University, Chicago, Illinois; ‡Department of Psychology, Fayetteville State University, Fayetteville, North Carolina; §Pain and Rehabilitation Clinic of Chicago, Chicago, Illinois, USA

ABSTRACT

Objective. To examine the taxonomic categories derived from a reorganization of the popular and useful Multidimensional Pain Inventory (MPI)/Multiaxial Assessment of Pain (MAP) system.

Design. Cluster-analytic procedures were utilized to analyze data from 976 pain patients who presented for treatment at a pain clinic. Ward’s method was utilized followed by seeded and unseeded k-means procedures and a discriminant function analysis.

Results. Three clusters were interpreted as: Interpersonally Focused, Stoic, and Adaptively Focused. Discriminant analyses followed by internal and external classification analyses provide evidence for the stability and replicability of the new clusters. Fisher’s linear discriminant classification function coefficients were produced for clinical use.

Conclusions. Meaningful differences exist between the new MPI clusters and those previously reported. The results provide a contribution toward improvements on the original and popular MPI/MAP system.

Key Words. Taxonomy; Chronic Pain; Cluster Analysis; MPI

Introduction

The classification of patients into meaningful groups is fundamental to medicine and psychology. Classifying patients with chronic pain is an ongoing challenge. Although efforts at classification have been productive, the usefulness, reliability, and validity of many attempts have been limited [1]. The existence of meaningful categories of chronic pain patients may indicate the need for different treatment modalities relative to membership [1–8]. Additionally, without meaningful categories to assess outcomes, aggregate treatment outcomes can dilute potentially significant results for any one group [1–8]. Although progress is being made, there is still much to be accomplished in the process of defining and refining a useful, reliable, and valid taxonomic system.

The 61-item Multidimensional Pain Inventory (MPI), a widely used instrument to assess pain patients, is often used in conjunction with the DOS-based Multiaxial Assessment of Pain (MAP) program to classify pain patients [5,6,9,10]. Given general patient acceptance and brevity of the MPI, its economic attractiveness, and its potential for predicting outcomes, the MPI/MAP system is probably the most popular and useful psychometric classification system for pain patients to date. The MPI/MAP system produces three categories of patients labeled as: Dysfunctional, Interpersonally Distressed, and Adaptive Coper/Minimizer [8]. The utility of this taxonomy is enhanced when used in conjunction with the Variable Responding (VR) scale to detect invalid profiles [11]. As there are nearly two decades of research to inform a revision, improving the instrument and the system would enhance its accuracy, reliability, validity, and clinical utility and help clinicians to make more
informed, reliable, and useful patient classifications to guide conceptualization and treatment.

There is mixed evidence for the generalizability and reliability of the MPI/MAP taxonomic categories. Several studies have replicated the results of the original cluster analyses with various groups of pain patients, that is, chronic low back pain, headache, temporomandibular disorders, and cancer [9,12–16], whereas others have not been able to replicate these results [17,18]. This suggests that perhaps the MPI/MAP system does not produce a full spectrum of categories or perhaps the categories are not stable, reliable, and replicable. In fact, a recent test–retest procedure over a 10-month period resulted in the MPI/MAP system changing the classification of 85% of the patients at least once [19]. Nonetheless, one of the challenges in evaluating any classification system is to impose an appropriate number and type of evaluative criteria. The process of evaluating cluster analyses, the mathematical process that has produced these categories, although based on empiric evidence, is often qualitative in nature. This evaluative process would benefit from reference to a consensus of quantitatively based minimum criteria.

Some of the concerns about the MPI/MAP system are related to its psychometric development. The system was developed from: 1) the 52-item West Haven-Yale MPI (WHYMPI) [20], not the 61-item instrument currently in use; 2) two relatively small samples of chronic pain patients (N = 122 and N = 100); and 3) only nine of the original 13 scales (a composite General Activity scale in place of the four activity scales). Additionally, the cluster analysis used to develop the categories utilized cluster means from the first sample to classify the second sample, meaning that the cluster solutions from the two samples were not independently generated [8]. These results generated decision rules that were used to create the DOS-based MAP program [5,8]. However, use of the MAP program produced a significant number (5%) of a fourth group, Hybrid/Anomalous, previously unaccounted for in the original cluster analysis [21].

Because the MAP/MPI system was derived from a cluster analysis of the scales from the WHYMPI, the stability of the system is linked to the original factor analysis of the WHYMPI [8,20]. Several studies have attempted to replicate the factor structure of the WHYMPI with mixed results [22–24]. These results are probably related to a number of issues. 1) Nine items were added to the MPI after the original factor analysis, the addition of which undoubtedly has had some effect [20]. 2) An exploratory factor analysis (EFA) of all 52 (or later 61) items was not performed. Instead of an EFA, a confirmatory factor analysis (CFA) was applied to Part I to verify the five theoretically derived scales while EFAs were used to derive the scales for Parts II and III [20]. 3) The sample size (N = 120) in the original factor analysis did not meet the suggested participants-to-variables ratio of at least 5 to 1 [20,25]. Consequently, the scales currently utilized in the MAP/MPI system do not reflect the actual factor structure of the MPI [23].

In 2001, Deisinger et al. addressed these issues by examining the factor structure of the 61-item version with large sample sizes (N = 452, N = 267, N = 253) and fairly stringent factor-analytic procedures [23]. An item-level EFA on the first sample included all 61 MPI items and a new factor structure was generated. The new structure was verified and cross-validated by performing CFAs on two additional independent samples. Based on these results, a revised scaling structure with three factors, nine scales, and 49 items was produced [23]. Although much of the scaling structure remained unchanged, there were a few important revisions. 1) There was a lack of support for two scales (Affective Distress and Life Control). These scales were eliminated. 2) Items from the Social Activities and Activities Away from Home scales were strongly related. These two scales were combined to form a single Recreation scale. 3) There was poor support for maintaining several items and these items were eliminated (see Table 1). Although still less than ideal, the reorganization

| Table 1 Item composition of the revised MPI scales (Deisinger et al. 2001 [23], p. 38) |
|---|---|---|
| Factor | Scales | Original MPI Section Item Number |
| Factor 1 Suffering | Interference | I: 2, 3, 4, 10, 11, 12, 18, 19, 23, 25, 27 |
| | Punishing Responses | II: 4, 7, 10 |
| | Pain Severity | I: 1, 7, 8, 16 |
| Factor 2 Social Support | Support | I: 5, 13, 20 |
| | Solicitous Responses | II: 5, 8, 11, 13, 14 |
| | Distracting Responses | II: 6, 9, 12 |
| Factor 3 Activity | Recreation | III: 3, 4, 7, 8, 11, 12, 15, 16, 19 |
| | Household Chores | III: 1, 5, 9, 13, 17 |
| | Outdoor Work | III: 2, 6, 10, 14, 18 |
provided by Deisinger et al. produced a more psychometrically sound instrument.

Previous methods for generating the current MPI/MAP categorical system have used less than optimal methods including only one cluster-analytic procedures, one evaluative index, and a lack of independent replication [8,9,12–16]. Although Burns et al. utilized two procedures (Ward’s method followed by the k-means procedure), the means from the Ward’s method solution were used to seed the k-means procedure and the results from the two methods cannot be considered independent [17]. Burns et al. also used two indices (clinical interpretability and the agglomeration schedule) to determine the optimal number of clusters, which is an improvement over a singular method, but less than ideal [17]. Finally, although Burns et al. improved on previous methods, the results were not replicated on a second sample.

Given the support, success, and limitations of the MAP/MPI system, investigation of the taxonomic categories derived from a cluster analysis of this reorganization was deemed worthwhile. This study seeks to investigate the taxonomic classification system generated from the Deisinger et al. revised MPI scoring and scales. This study also seeks to exemplify a thorough, sound, and effective cluster-analytic procedure. A well-designed cluster-analytic study includes the examination of the stability of a cluster solution by independently applying more than one clustering procedure [26]. Using Ward’s method as the first procedure and the k-means as the second procedure is recommended and frequently utilized [27,28]. Additionally, the optimal number of clusters ideally is determined by logical and careful evaluation of a number of heuristic and statistical indices that may vary across steps [26,29,30]. Finally, the cluster solution should be replicated on a second, independent sample [26].

This study will derive and evaluate a taxonomic classification system generated from the revised scales of the MPI using recommended methods in a step-by-step manner [23]. A summary of the steps utilized in the analysis is provided in Table 2.

Method

Participants

The initial data set was comprised of archival data from a clinical sample of 1,573 consecutive new patients from an outpatient pain rehabilitation clinic and hospital in the Chicago area. Initial inclusion criteria were administration of the MPI as well as complete participant information for sex, age, duration of pain, and International Association for the Study of Pain (IASP) site.

Table 2 Steps for data analysis

<table>
<thead>
<tr>
<th>Process</th>
<th>Activities</th>
<th>Steps</th>
</tr>
</thead>
</table>
| Data preparation                             | 1. Apply inclusion criteria  
|                                              | a. MPI administered  
|                                              | b. Complete information for sex, age, duration of pain, and IASP site  
|                                              | c. VR scale of 15 or below  
|                                              | 2. Calculate revised MPI scales for each participant (21)  
|                                              | 3. Randomly split data into two equivalent samples  
|                                              | 4. Assess for significant differences between S1 and S2  
| Identify the optimal number of clusters      | 1. Apply Ward’s method to S1  
| and the optimal cluster procedure            | 2. Determine the most probable cluster solutions from Ward’s method using a number of indices  
|                                              | 3. Apply k-means iterative partitioning procedures to S1  
|                                              | 4. Determine the optimal cluster solution using a number of indices  
|                                              | 5. Determine the optimal clustering procedure for obtaining this solution  
| Replicate cluster solution across data sets  | 1. Apply optimal clustering procedure to S2  
|                                              | 2. Compare results from S1 and S2 with various measures  
| Examine solution with variables not used     | 1. Apply optimal clustering procedure to entire sample (N = 976)  
| used to generate the cluster solution        | 2. Conduct ANOVAs testing for differences in mean BDI, STAI-T, and MPQ-Tot scores  
| Assess clinical interpretability             | 1. Examine and evaluate profile patterns using scales, BDI, STAI-T, and MPQ-Tot  
| Initial discriminant analysis                | 1. Perform initial discriminant analysis on S1  
|                                              | 2. Perform internal classification analysis  
|                                              | 3. Perform external classification analysis by applying results of S1 to S2  
| Final discriminant analyses                  | 1. Perform discriminant analysis on all 976 cases  
|                                              | 2. Report Fisher’s linear discriminant function coefficients with directions for use  

MPI = Multidimensional Pain Inventory; BDI = Beck Depression Inventory; STAI-T = State-Trait Anxiety Inventory—Trait Scale; SF-MPQ-Tot = Short Form-McGill Pain Inventory—Total Scale.
A Taxonomy of Chronic Pain Patients

Table 3 Distribution of International Association for the Study of Pain (IASP) pain sites in sample

<table>
<thead>
<tr>
<th>Site</th>
<th>Percent of Total (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head, face, mouth</td>
<td>3.4 (33)</td>
</tr>
<tr>
<td>Cervical</td>
<td>8.8 (86)</td>
</tr>
<tr>
<td>Upper shoulder, upper limbs</td>
<td>11.5 (112)</td>
</tr>
<tr>
<td>Thoracic</td>
<td>3.4 (33)</td>
</tr>
<tr>
<td>Abdominal</td>
<td>3.4 (33)</td>
</tr>
<tr>
<td>Low back, lumbar spine, sacrum coccyx</td>
<td>44 (429)</td>
</tr>
<tr>
<td>Lower limbs</td>
<td>24 (234)</td>
</tr>
<tr>
<td>Pelvic</td>
<td>&lt;1 (8)</td>
</tr>
<tr>
<td>Anal, perineal, genital</td>
<td>&lt;1 (6)</td>
</tr>
<tr>
<td>More than three sites</td>
<td>&lt;1 (2)</td>
</tr>
</tbody>
</table>

The revised scaling structure of the MPI (described earlier) includes three factors, nine scales, and 49 of the original items [23]. Each item is rated on an anchored seven-point frequency scale ranging from 0 = never to 6 = very often. The mean score for each of the nine scales was determined by summing the scores for all items in the respective scale and dividing by the number of items in that scale. Scale scores range from 0 to 6, with higher scores indicating a greater frequency of occurrence. The internal consistency of the revised scales (Cronbach's alpha) ranges from 0.74 to 0.90 [23]. Participants were administered the 61 item MPI. Data were entered using the MAP program and exported to an American Standard Code for Information Interchange (ASCII) data file.

Items were then extracted and reorganized into the revised scaling structure according to recommendations of Deisinger et al. (see Table 1) [23].

Beck Depression Inventory
The Beck Depression Inventory (BDI) is a self-report measure widely used to assess depression in an extensive variety of patients. The 21 items assess depressive symptoms, with each item rated on an anchored four-point scale ranging from 0 to 3. Scores range from 0 to 63 with higher scores indicating greater levels of depression. The BDI has demonstrated satisfactory internal consistency (0.86) and reliability. Test–retest coefficients range from 0.48 to 0.86 depending on retesting interval and population. Content, concurrent, and discriminant validity are generally favorable. High correlations (0.72–0.76) are found between the BDI and other self-report measures as well as clinician ratings of depression [31,32].

State-Trait Anxiety Inventory-Trait Scale
The State-Trait Anxiety Inventory (STAI) is comprised of two self-report scales measuring state (STAI-S) and trait (STAI-T) anxiety. Only the STAI-T scale was used in this study. The STAI-T assesses the relatively stable general emotional condition characterized by subjective, consciously perceived feelings of tension, apprehension, nervousness, worry, and heightened activation of the autonomic nervous system. These trait features are thought to remain relatively stable despite levels of situational stress. Each of the 20 STAI-T items is rated on an anchored four-point frequency scale ranging from 1 = almost never to 4 = almost always. Scores range from 20 to 80 with higher scores indicating greater levels of anxiety. The internal consistency reliabilities measured by Cronbach's alpha are high (0.90). Test–retest reliability has tended to be relatively high for the Trait scale and low for the State scale. Construct and concurrent validity of the STAI have been satisfactorily demonstrated in a number of populations [21,33,34].

Short Form-McGill Pain Questionnaire
The Short Form-McGill Pain Questionnaire (SF-MPQ) is a 15-item adjective checklist of pain descriptors (11 sensory; 4 affective) rated on a four-point intensity scale from 0 = none to 3 = severe. The summary measure MPQ—Total (MPQ-Tot) was utilized in this study. The MPQ-Tot provides a score ranging from 0 to 45, with higher scores indicating greater pain intensity. Correlation coefficients between the SF-MPQ-
Tot score and the Long Form-MPQ range from 0.77 to 0.92. Similar to the Long Form, the SF-MPQ is sensitive to the effects of traditional therapies and appears capable of providing discriminatory information about different pain syndromes [35,36].

Procedure
This study was submitted and approved by all appropriate Institutional Review Boards at participating institutions. Patients were administered the assessment instruments as part of routine clinical practice. No incentives were provided. Scores, medical diagnoses, and demographic information were entered into a computerized record system and then exported to a database. No information was maintained in the database that would allow the identification of individuals. SPSS version 9 (SPSS Inc., Chicago, IL, USA) was used for all analyses.

The MPI items were reorganized into the revised scaling structure [23]. A random split procedure generated two demographically equivalent samples: Sample 1 (S1; N = 489) and Sample 2 (S2; N = 487). Tests of significance demonstrated no significant differences between the samples with regard to sex, Chi-square (1, N = 976) = 0.006, \( P = 0.94 \); age, \( F_{1,975} = 1.7, P = 0.20 \); duration of pain, E1,975 = 1.8, \( P = 0.18 \); and IASP site, Chi-square (8, N = 976) = 0.31, \( P = 1.00 \).

A summary of the steps utilized in the analysis is provided in Table 2.

Step 1
Method
The patient taxonomic classification method began by applying Ward's method to hierarchically “cluster” S1 patients into groups based on similar MPI score patterns or “profiles” [29,37]. An optimal solution shows maximal similarity of patient profiles within clusters and maximal dissimilarity of patient profiles between clusters. Four indices were used to assess the cluster solutions produced by Ward’s method. 1) The agglomeration schedule: a large difference in sequential coefficients immediately after two clusters are merged is optimal [28]. 2) Clear visual separation of clusters in the dendrogram is optimal. 3) High pseudo-\( F \) values (a measure of agreement between the mean scale scores and the cluster group assignment) are optimal [38]. 4) Significant differences between mean scales scores among clusters within a solution are Optimal [29]. For instance, in a three-cluster solution, one would expect the value of the Pain Severity scale to be significant different in clusters 1, 2, and 3. Data files of the scale means for the optimal solutions produced by Ward’s method were created to seed the k-means cluster analysis for Study 2.

Results
Examination of the agglomeration schedule did not reveal large differences in sequential coefficients. The dendrogram and the pseudo-\( F \) values for solutions with two through 10 clusters revealed support for the two (pseudo-\( F = 97.15 \)), three (pseudo-\( F = 91.89 \)), and four (pseudo-\( F = 79.18 \)) cluster solutions. The two cluster solution yielded nonsignificant differences in three out of the nine comparisons. The three-cluster solution yielded nonsignificant differences in three out of 18 comparisons. The four-cluster solution yielded nonsignificant differences in 14 out of 54 comparisons. The patient profiles grouped into the three-cluster solution were determined to be maximally dissimilar.

Step 2
Method
A high degree of similarity between clusterings of the same data using different methods indicates a high degree of profile stability [39]. The stability of the cluster solutions with the highest pseudo-\( F \)s from Step 1 (the two-, three-, and four-cluster solutions) was evaluated with the k-means iterative clustering procedure. The k-means procedure begins with a priori assumptions about initial cluster centers [27]. Initial cluster centers are sometimes specified beforehand or “seeded,” but if not specified, the procedure will seed itself by selecting the first few cases as the initial cluster centers [27,28]. Although the k-means procedure performs best when well seeded, using the results from Step 1 to seed the procedure would not provide independent results. Both seeded and unseeded procedures were applied to S1.

The stability of the solutions generated by the k-means procedure was evaluated with four criteria: 1) comparison of the results from the seeded and the unseeded k-means applications; 2) the pseudo-\( F \) criterion; 3) Cohen’s kappa; and 4) significant differences between scale means among clusters within each solution. Cohen’s kappa, likely to control for chance agreement, is often utilized as a measure of agreement between two different classifications of the same data set, in this case Ward’s method and the k-means cluster assignments. An agreement kappa will range from 1.0
A Taxonomy of Chronic Pain Patients

perfect agreement) to 0.0 (no agreement), with a kappa value of 0.70 indicating a moderate to high degree of agreement [40].

The optimal method for replicating the solution was then determined. The optimal method was expected to produce significant differences among cluster within solutions, adhere to the psychometric qualities of the revised MPI, and appear clinically meaningful.

Results

The three-cluster solutions produced identical final cluster centers and cluster membership in the seeded and unseeded applications, a high pseudo-$F$ value ($F = 110.9$), the fewest number of non-significant differences among the mean scale scores, the highest Cohen’s kappa coefficient ($kappa = 0.59$), and appeared optimal.

The k-means procedure was determined to be the optimal method for replicating the results of S1 with S2. This procedure produced only one nonsignificant difference in scale means in the three-cluster solution (Pain Severity) while Ward’s method produced three nonsignificant differences. The pseudo-$F$ values were greater for the k-means for all solutions. The k-means solution also appeared clinically meaningful and similar to the original MPI cluster analysis results [8].

Step 3

Method

Because the accuracy of a cluster solution with an unknown “true” classification cannot be directly measured, accuracy is often assessed by replicating a cluster solution across data sets [39]. Although replicability is not strong evidence for validity, the failure of a cluster solution to replicate is reason for rejecting the solution. Using the results of Step 2, the k-means procedure was applied to S2 using the final cluster means from the k-means three-cluster solution as “seeds” for the procedure.

Four criteria were used to assess the cluster solutions across data sets: 1) Pearson’s product moment correlations; 2) visual inspection of graphs of mean scale scores; 3) cluster membership frequency; and 4) significant differences in scale means among clusters within each solution. Pearson product moment correlations are often used to assess similarity of “profiles” produced by psychological instruments, that is, MMPI [29]. An agreement correlation coefficient will range from $-1$ to 1 with 1 indicating a perfect positive relationship, $-1$ indicating a perfect negative relationship, and 0 indicating no linear relationship between profiles.

Results

Comparison of the results from S1 and S2 indicated that the three-cluster solution generated by the k-means procedure was stable across data sets. The Pearson product moment correlations between S1 and S2 cluster 1, 2, and 3 profiles were each $r^2 = 0.99, P < 0.01$ (two-tailed). Visual inspection of the graphs revealed considerable similarity. Membership frequency was also consistent across data sets. S1/S2 frequency ratios were cluster 1: 35%/34%; cluster 2: 21%/24%, cluster 3: 44%/42%. There was also a similar pattern of significant differences among scales. Both S1 and S2 produced nonsignificant differences between clusters 2 and 3 on Pain Severity. However, S2 alone produced a nonsignificant difference between clusters 1 and 3 on Punishing Responses. The evidence supported this procedure and solution to be largely stable across data sets, indicating a high degree of accuracy and providing evidence for validity.

Step 4

Method

Validity of a cluster solution is often evaluated by comparing the clusters on variables not used to generate the solution. The k-means procedure was applied to the original group of 976 participants, using the final cluster centers from the optimal cluster solution in Study 2 as the “seeds.” The resulting cluster assignment for each participant was recorded as the final cluster assignment for that case. Using these results, tests of significance were used to examine clusters 1, 2 and 3 with regard to the mean BDI, STAI-T, and SF-MPQ-Tot scores for each cluster. The final cluster assignment was then cross-tabulated with the original MAP assigned categories. Finally, clinical interpretability of the clusters was considered and the clusters were labeled.

Results

After the final cluster assignment for each of the 976 cases was calculated, significant differences were found among all the scaled scores in each of the three clusters except between cluster 2 and cluster 3 on the Pain Severity scale.

Evaluation of the final cluster solution on variables not used to generate the solution indicated adequate validity. Tests of significance revealed significant differences among the clusters for all
three variables (BDI, $F_{2,341} = 13.0, P < 0.01$; STAI-T, $F_{2,347} = 17.4, P < 0.01$; SF-MPQ-Tot, $F_{2,353} = 18.2, P < 0.01$). Post hoc tests revealed significant differences in BDI scores between clusters 1 and 3, in STAI-T scores between clusters 1 and 3, and 2 and 3, and in SF-MPQ-Tot scores between clusters 1 and 2, and 1 and 3 (see Table 4).

Examination of the profile patterns for each cluster in conjunction with the BDI, STAI-T, and SF-MPQ-Tot data revealed distinct profiles with particular characteristics (see Figure 1 and Tables 4–6). The clusters were labeled as follows.

Cluster 1: The Interpersonally Focused Profile (35%).

The hallmark of this profile is high scores on all scales in Factor 2/Social Support combined with high and/or intermediate scores in Factor 1/Suffering scales and low scores on all scales in Factor 3/Activity. Participants with this profile report significant others are highly supportive, helpful, worried, and attentive. They report that others often take over their jobs, try to get them to rest, obtain pain medication for them, or help distract them from pain, and that pain has significantly changed relationships with significant others. They also report that occasionally their significant others get

<table>
<thead>
<tr>
<th>Measure</th>
<th>Cluster 1 (Mean, SD)</th>
<th>Cluster 2 (Mean, SD)</th>
<th>Cluster 3 (Mean, SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDI</td>
<td>1 (13.7, 8.4)</td>
<td>2 (11.3, 8.5)</td>
<td>3 (8.9, 6.5)**</td>
</tr>
<tr>
<td></td>
<td>2 (11.3, 8.5)</td>
<td>3 (8.9, 6.5)</td>
<td></td>
</tr>
<tr>
<td>STAI-T</td>
<td>1 (43.0, 14.5)</td>
<td>2 (40.2, 12.6)</td>
<td>3 (34.5, 9.6)**</td>
</tr>
<tr>
<td></td>
<td>2 (40.2, 12.6)</td>
<td>3 (34.5, 9.6)**</td>
<td></td>
</tr>
<tr>
<td>SF-MPQ-Tot</td>
<td>1 (25.8, 9.5)</td>
<td>2 (22.1, 9.2)**</td>
<td>3 (19.2, 8.0)**</td>
</tr>
<tr>
<td></td>
<td>2 (22.1, 9.6)</td>
<td>3 (19.2, 8.0)</td>
<td></td>
</tr>
</tbody>
</table>

** $P < 0.01$.

BDI = Beck Depression Inventory (BDI); STAI-T = State-Trait Anxiety Inventory—Trait Scale; SF-MPQ-Tot = Short Form-McGill Pain Questionnaire—Total Score.

Examination of the mean scale scores for each cluster revealed distinct scoring patterns. The scoring patterns were examined and interpreted as: 1) Interpersonally Focused, high scores on the Social Support scales (S, SR, DR) combined with high/intermediate scores on the Suffering scales (I, PR, PS), and low scores on Activity (R, HC, OW); 2) Stoic, high/intermediate scores on the Suffering scales (I, PR, PS) with particularly high scores on Punishing Responses (PR), low scores on the Social Support scales (S, SR, DR), and intermediate Activity scores (R, HC, OW); and 3) Adaptively Focused, high Activity scores (R, HC, OW) combined with low Suffering scores (I, PR, PS), and intermediate Social Support scores (S, SR, DR). See Step 4.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Cluster 1 (35.2%) Interpersonally Focused</th>
<th>Cluster 2 (23.7%) Stoic</th>
<th>Cluster 3 (41.1%) Adaptively Focused</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1 Suffering</td>
<td>High/intermediate</td>
<td>High/intermediate</td>
<td>Low</td>
</tr>
<tr>
<td>Factor 2 Social Support</td>
<td>High</td>
<td>Low</td>
<td>Intermediate</td>
</tr>
<tr>
<td>Factor 3 Activity</td>
<td>Low</td>
<td>Intermediate (mean = 11.3)</td>
<td>High</td>
</tr>
<tr>
<td>BDI</td>
<td>High (mean = 13.7)*</td>
<td>Intermediate (mean = 43)*</td>
<td>Low (mean = 8.9)*</td>
</tr>
<tr>
<td>STAI-T</td>
<td>Intermediate (mean = 40.2)*</td>
<td>Low (mean = 34.5)*</td>
<td></td>
</tr>
<tr>
<td>SF-MPQ-Tot</td>
<td>Intermediate (mean = 22.1)*</td>
<td>Low (mean = 19.2)*</td>
<td></td>
</tr>
</tbody>
</table>

Means with the same superscript are significantly different, $P < 0.01$.

BDI = Beck Depression Inventory; STAI-T = State-Trait Anxiety Inventory—Trait Scale; SF-MPQ-Tot = Short Form-McGill Pain Questionnaire—Total Score.

![Figure 1](image-url)
angry, irritated, or frustrated with them. They report intermediate to high levels of pain and that pain has interfered greatly in their ability to work, play, and engage in most all activities.

This profile roughly corresponds to the original Dysfunctional profile [8]. Cross-tabulation indicates a moderate correspondence with 78% of Interpersonally Focused profiles also categorized as Dysfunctional (268/344; see Table 6). Certain scales that defined the Dysfunctional profile, however, Affective Distress and Life Control, no longer exist in the revised scaling structure, precluding a straightforward comparison. In contrast to the Dysfunctional profile, which focuses on affective distress and loss of control, the Interpersonally Focused profile is characterized by an “other focus” or a “relationship focus” as reflected by the dominance of the Factor 2/Social Support scales. This, in addition to the relative lack of focus on adaptive functioning (Factor 3/Activity), suggests that the participants with this profile tend to focus on how others can and do meet their needs to the exclusion of how they might continue to perform the activities in which they normally engage.

The term Interpersonally Focused was determined to provide an accurate description of this group and have fewer negative connotations than the term Dysfunctional. Consistent with a behavioral model of chronic pain, high levels of solicitous and negative significant other behaviors appear to be associated with increased dysfunction and pain for these participants [41]. Participants with this profile reported intermediate to high levels of pain, the lowest levels of adaptive functioning, and highest BDI, STAI-T, and SF-MPQ-Tot scores (see Table 5).

Cluster 2: The Stoic Profile (24%). The hallmark of this profile is high to intermediate Factor 1/Suffering scores with particularly high scores on Punishing Responses, low scores in Factor 2/Social Support, and intermediate Factor 3/Activity scores. Participants with this profile describe significant others as offering little assistance as well as frequently getting angry, frustrated, or irritated with others. They report that pain has interfered significantly in their ability to work, play, and engage in most activities, but they appear to maintain an intermediate level of activity despite this interference.

This profile roughly corresponds to the original Interpersonally Distressed profile. Cross-tabulation shows 54% of Stoic profiles categorized as Interpersonally distressed (125/231; see Table 6). Again, the differences in scale organization confound a straightforward comparison, that is, the Distress scale was eliminated. However, both the Interpersonally Distressed and the Stoic profiles are characterized by low Social Support scores and high scores on Punishing Responses. Interestingly, these participants report levels of suffering roughly equivalent to those in cluster 1/Interpersonally Focused. Despite high levels of suffering, the participants with this profile maintained intermediate levels of activity and intermediate scores on the BDI, STAI-T, and SF-MPQ-Tot (see Table 5). The term Stoic was deemed descriptive of this group and to have fewer negative connotations than Interpersonally Distressed.

Cluster 3: Adaptively Focused (41%). The hallmark of this profile is high Factor 3/Activity scores combined with low Factor 1/Suffering scores and intermediate Factor 2/Social Support scores. Interestingly, participants with this profile do not report significantly more or less pain than the Stoic profile. Despite significant pain, moderate social support, and moderate pain interference, participants with this profile report that they frequently engage in a number of daily activities and appear focused on their adaptive functioning. Participants with this profile were associated with the lowest BDI, STAI, and SF-MPQ-Tot scores.

This profile roughly corresponds to the original Adaptive Coper/Minimizer profile. Cross-tabula-
tion shows 43% of Adaptively Focused profiles categorized as Adaptive Copers. As Turk and Rudy [8] suggest, these more adaptive participants may be denying or minimizing the extent of their pain problems or they may be engaged in more effective coping strategies. Given the limitations of the survey instrument, actual adaptiveness is impossible to discern. The term Adaptively Focused was determined to be both descriptive and accurate as it describes both minimizers and adaptive copers as focused on maintaining high levels of adaptive functioning.

Step 5
Method
Because there are significant differences in the item content and scale organization of the MPI/MAP system and the revised MPI scales, a quantitative comparison of results must be viewed with caution. Nevertheless, as might be expected, there is considerable agreement between the cluster assignments of both systems (see Table 6). However, cross-tabulation of the final cluster assignments from Study 4 with the corresponding original MAP cluster assignments raised a number of questions. Although there is considerable similarity, a substantial portion of patients originally classified as Dysfunctional were classified as Adaptively Focused (N = 102, 25%) or Stoic (N = 36, 9%). This merited an exploratory investigation within the Dysfunctional group with the expectation that the Interpersonally Focused group would demonstrate significantly higher scores than the other two groups on the BDI, STAI-T, and/or the SF-MPQ-Tot. Likewise, within the Interpersonally Distressed group, the Stoic group was expected to demonstrate significantly higher scores than the Adaptively Focused group, and within the Adaptive Coper group, the Interpersonally Focused and the Stoic groups were expected to demonstrate significantly higher scores than the Adaptively Focused group. One-way ANOVAs followed by post hoc testing with Fisher’s Least Significant Difference (LSD) were used to test for significant differences in mean BDI, STAI-T, and SF-MPQ-Tot scores among the Interpersonally Focused, Stoic, and Adaptively Focused groups within the original MPI clusters/groups.

Results
A significant difference in mean SF-MPQ-Tot scores was found among the three new profiles within the original Dysfunctional group, $F_{2,147} = 3.72, P < 0.03$. Post hoc testing revealed significantly different SF-MPQ-Tot scores between the Interpersonally Focused (mean = 25.59) and the Stoic (mean = 16.92) groups and the Interpersonally Focused (mean = 25.59) and the Adaptively Focused (mean = 20.85) groups. This indicates that within the original Dysfunctional group, the Interpersonally Focused group indeed differs from the other two groups in the expected direction on a measure not used to generate either cluster solution. The Interpersonally Focused group reported significantly more pain than the Stoic and the Adaptively Focused groups.

A significant difference in mean STAI-T scores was found among the three new profiles within the original Interpersonally Distressed group, $F_{1,58} = 6.18, P < 0.02$. Post hoc testing revealed significantly different STAI-T scores between the Stoic (mean = 37.17) and the Adaptively Focused (mean = 28.35) groups. This indicates that within the original Interpersonally Distressed group, the Stoic and the Adaptively Focused groups indeed differ from one another in the expected direction on a measure not used to generate either cluster solution. The Stoic group reported significantly more anxiety than the Adaptively Focused group.

No significant differences were found among the three new profiles within the original Adaptive Coper/Minimizer group.

Step 6
Method
The samples identified as S1 and S2 were utilized for the initial discriminant analysis and the internal and external classification analysis [42]. The initial discriminant analysis was performed on S1. The final cluster assignment from Study 4 was used as the original classification criterion. Mean scale scores were entered as predictive variables in a stepwise procedure with the minimum significance of $F$ to enter as 0.05 and the minimum significance of $F$ to remove as 0.10. At each step, the variable that minimized the overall Wilks’ lambda was entered. Two discriminant function formulas were utilized to assign the case to the cluster with the highest probability of membership [28]. As an internal classification analysis, cases classified in this manner were then compared with the original classification criterion. The percentage of patients consistently reclassified in S1 was then calculated.

An external classification analysis was then performed by applying the results of S1 to S2 using a method of classification that a clinician in the field would be able to easily apply to MPI
A Taxonomy of Chronic Pain Patients

321

Results scored with the revised scoring method. Fisher's discriminant functions, a set of coefficients generated by the discriminant analysis on S1, were used to classify S2 [43,44]. See Table 7 for the procedure. The external classification analysis was completed by calculating the percentage of patients consistently reclassified and comparing S2’s predicted cluster membership with S2’s final cluster assignment using Cohen’s kappa.

**Results**

When the discriminant analysis was performed on S1, seven of the nine scales produced a significant decrease in Wilks’ lambda. Pain Severity and Outdoor Work did not contribute to a significant decrease in Wilks’ lambda in S1 and were not included in the analysis any further. Tests of the function formulas 1 though 2 were positive, producing a Wilks’ lambda of 0.17 and Chi-square (18) = 1692.93, P < 0.001, indicating a relatively low proportion of total variance in discriminant scores not explained by differences among the clusters. The discriminant function formulas reclassified 94.3% of the 976 cases consistent with the final cluster assignment from Step 4. Cohen’s kappa, used to compare the predicted and final cluster memberships in all 976 cases, was significant (kappa = 0.91, P < 0.01). The final Fisher’s linear discriminant function coefficients and directions for clinical application are reported in Table 7. Using the procedure delineated in Table 7, clinicians can classify any set of MPI scales scored with the Deisinger revision.

**Discussion**

This investigation examined the taxonomic categories derived from the Deisinger et al. reorganization utilizing an effective cluster-analytic procedure and represents a contribution toward improvements on the original MPI/MAP system [23]. The results are derived from replicable, cross-validated, revised scaling structure of the MPI, an improvement over the scaling structure used in the MPI/MAP system. The samples utilized in this study were larger (N = 489 vs N = 122) and more representative than those used in the original analysis [8]. The methodology in

**Table 7** Final Fisher’s linear discriminant classification function coefficients*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interpersonally Focused</td>
<td>4.188</td>
<td>3.967</td>
<td>3.528</td>
</tr>
<tr>
<td>Punishing responses</td>
<td>-2.508</td>
<td>0.727</td>
<td>2.703</td>
</tr>
<tr>
<td>Pain severity</td>
<td>2.987</td>
<td>2.788</td>
<td>2.59</td>
</tr>
<tr>
<td>Support</td>
<td>3.621</td>
<td>1.9</td>
<td>3.777</td>
</tr>
<tr>
<td>Solicitous responses</td>
<td>1.565</td>
<td>0.226</td>
<td>0.774</td>
</tr>
<tr>
<td>Distracting responses</td>
<td>0.884</td>
<td>-0.138</td>
<td>0.302</td>
</tr>
<tr>
<td>Recreation</td>
<td>2.543</td>
<td>2.976</td>
<td>3.268</td>
</tr>
<tr>
<td>Household chores</td>
<td>0.33</td>
<td>0.875</td>
<td>1.551</td>
</tr>
<tr>
<td>Outdoor work</td>
<td>0.514</td>
<td>0.742</td>
<td>0.869</td>
</tr>
<tr>
<td>Constant</td>
<td>-36.879</td>
<td>-25.047</td>
<td>-32.812</td>
</tr>
</tbody>
</table>

* To classify a case into a cluster, a) begin with cluster 1, b) multiply each of the scales by its respective classification coefficient, c) add the products to the respective constant to obtain a single value, d) repeat steps a and b with clusters 2 and 3 to obtain three values corresponding to the three clusters. Classify the case into the cluster that produced the highest of the three values.

**Step 7**

**Method**

The results from a discriminant analysis using a larger data set is likely to be more generalizable than results from smaller data sets. The discriminant analysis was repeated utilizing the entire data set (N = 976). The final cluster assignment from Step 4 was used as the classification criterion. Mean scale scores were entered as predictive variables in a stepwise procedure with the minimum significance of F to enter as 0.05 and the minimum significance of F to remove as 0.10. At each step, the variable that minimized the overall Wilks’ lambda was entered. (See Study 6 for more details.) As a final internal classification analysis, the results of this classification procedure were compared with the original classification criterion using Cohen’s kappa.

**Results**

When the discriminant analysis was repeated on all 976 cases, all nine scales produced a significant decrease in Wilks’ lambda. Tests of the function formulas 1 though 2 were positive, producing a Wilks’ lambda of 0.17 and Chi-square (18) = 1692.93, P < 0.001, indicating a relatively low proportion of total variance in discriminant scores not explained by differences among the clusters. The discriminant function formulas reclassified 94.3% of the 976 cases consistent with the final cluster assignment from Step 4. Cohen’s kappa, used to compare the predicted and final cluster memberships in all 976 cases, was significant (kappa = 0.91, P < 0.01). The final Fisher’s linear discriminant function coefficients and directions for clinical application are reported in Table 7. Using the procedure delineated in Table 7, clinicians can classify any set of MPI scales scored with the Deisinger revision.
this study adhered more closely to currently accepted recommendations for clustering procedures, an improvement over the methods used in the origination of the MPI/MAP system and the studies that replicated these findings. As previously noted, the stability of a cluster solution is assumed to be directly related to the replicability of the cluster solution across methods and data sets [39]. This analysis independently applied Ward’s method and the k-means procedures to S1. Through consideration of a number of statistical and heuristic indices, a three-cluster solution generated by the k-means procedure was determined to be optimal. This solution generated the highest kappa coefficients, was identical across both the seeded and unseeded k-means applications, provided for adequate pseudo-$F$ values, and yielded the least number of nonsignificant differences among mean clusters. This solution was then replicated on another large sample, S2. Pearson product moment correlations between S1 and S2 for cluster 1, 2, and 3 in the three-cluster solution were highly significant. These findings indicate that the resultant cluster solution has a high degree of stability and replicability suggesting that clinicians will find the categories derived from the new scoring and classification system more stable and reliable than the old MPI/MAP system.

The clusters were also validated with measures not used to generate the solution. Tests of significance revealed significant differences among the clusters for the BDI, the STAI-T, and the SF-MPQ-Tot. Significant differences were found between at least two clusters for each of these variables, indicating that the clusters consist of groups with distinct features that can be assessed by other instruments and are not simply an artifact of the analysis.

Although there is value in comparing the differences in the original and the revised patient categories, note that these two sets of categories are inherently qualitatively different. Recall that the Deisinger et al. revision eliminated scales that helped define Dysfunctional profile (Affective Distress, Life Control). These scales do not define the corresponding Interpersonally Focused profile. Conceptually, the BDI and the STAI-T contain items that access affective distress, loss of control, and related factors. Consequently, these instruments will have a different association with the Dysfunctional profile than with the Interpersonally Focused profile. Although the results in Step 5 are of interest and demonstrate that distinct categories exist within the original clusters, comparisons must be viewed with this caveat in mind.

With the large sample sizes and the full range of scores within the scales, this investigation also succeeded in avoiding transformation of the data into standardized scores, allowing for a full and overt examination of scale means, increased clinical accessibility, and the potential for compelling clinical comparisons in practice. Reporting the mean scale scores allows clinicians who use the revised scoring/scaling structure to directly compare their patients’ scores with these results in their office. A clinician can then score a patient’s MPI responses according the Deisinger et al. revision, plot the scale scores, and directly compare the results to the profiles in Figure 1. Category membership can be calculated using the procedure described in Table 7.

Case Example: Mr. A presented to the clinic with chronic low back pain, onset subsequent to falling backward off a backyard deck during a social event. Mr. A reports experiencing a constant dull pain (6 on a scale of 0–10), with occasional sharp shooting pains radiating down the back of his left leg (9 on a scale of 0–10). The eventual diagnosis was chronic neuropathic pain secondary to nerve trauma related to injury. Mr. A has been unable to work since the accident and reports feelings of inadequacy, low motivation, and apathy. He reports that the pain interferes with most usual day-to-day activities such as walking for any distance, preparing snacks, yard work, watching TV, or reading. He reports that his spouse and his teenage daughter have been very supportive and patient with him. They reportedly get frustrated with him sometimes, but do not complain that he is unable to engage in family activities and continue his role as “breadwinner” and “head of household.” He reports that they do a great deal to help him manage his pain such as trying to get him to “take it easy” and “get my mind of it.” He reports that “they can really tell when he is having a hard time” and that he “looks forward all day to them coming home” from work and school. The clinician observed that during the intake, the daughter and the wife were highly concerned about Mr. A’s comfort and that Mr. A rarely even changed position without making eye contact with his wife.

After calculating the mean scale scores using the revised scoring system and applying the calculation in Table 7, it was determined that Mr. A was classified as Interpersonally Focused. These results in addition to data from the clinical interview indicate that Mr. A is overly focused on others ensuring that his needs are met which results in a decreased focus on developing methods to self-manage his pain and an overall increase in pain and a decrease in physical activity. The treatment plan was designed to include interventions to increase Mr. A’s pain self-management skills, decrease (but not eliminate) the level of solicitous behaviors performed by the family and family involvement in pain management, and slowly incorporate a number of physical activities of value to the household to be determined by Mr. A and his family as treatment progressed. The plan included periodic reassessment of Mr. A using the MPI and the revised scoring system with the objective of obtaining significant decreases in LPS, S, SR, and DR while increasing R, HC, and OW.

The original MPI results were converted to T-scores. Although useful, clinicians cannot directly
compare patients' raw scores with results reported in T-scores without a conversion scale, more information, and additional steps. Furthermore, the original sample was small (N = 122) and the likelihood that T-scores from small samples may contain artifact is high.

The meaning and the role of Social Support scales in the new as well as the original taxonomic system are worthy of further discussion. As would be expected, all the Social Support scales (Support, Solicitous Responses, and Distracting Responses) load heavily on a single Social Support factor in the Deisinger et al. study, with correlation coefficients ranging from 0.66 to 0.79 [23]. These scales also maintain the same correlational direction and function similarly in each cluster/category. Although this speaks to sound psychometric organization, one might not expect the three scales that comprise Social Support to be positively associated to this degree. In general, Social Support is considered a relatively positive influence while Solicitous Responses are generally considered a relatively negative influence. The Support, Solicitous Responses, and Distracting Responses scales in the original MPI/MAP system follow a similar trend [8].

This suggests a more complex relationship between significant other behavior and the behavior of chronic pain patients than has been highlighted in the past. Increased solicitous behaviors by significant others are often associated with increased pain, increased pain behaviors, and decreased activity while supportive, distracting, and encouraging responses are associated with decreased pain, decreased pain behaviors, and increased activity [41,45,46]. In contrast to the literature, in this investigation high Social Support scales are associated with increased pain and low levels of activity. Low Social Support scales are associated with moderate activity. Intermediate Social Support scales are associated with the highest level of activity. Examination of these results suggests a nonlinear relation between significant other behavior and the behavior of chronic pain patients, suggesting that the clinical presentation of social support must be examined carefully. While a certain amount of social support is desirable, it appears that optimal functioning is associated with a balance between support and independent functioning.

There are, perhaps, a number of underlying reasons for the role Factor 2/Social Support plays in these results. On face value, the benefits of a low level of solicitousness might be neutral-ized when combined with low levels of supportive and distracting behaviors and high levels of punishing responses. However, another explanation lies in the self-report nature of the instrument. Patient self-report does not always approximate an accurate measure of significant other solicitousness [47]. In fact, results based on patient report of significant other solicitousness are often not in accordance with the general finding that it is a relatively negative influence. Given this information, clinicians may look to add alternative sources of information to understand actual vs perceived significant other solicitousness.

Given the discrepancies between patient self-report and observed solicitous responses, the accuracy of patient self-report when assessing distracting, encouraging, and even supportive significant other behaviors comes into question. These issues may need to be addressed with validation studies or by more fully acknowledging the self-report nature of the Social Support scales when interpreting results. Clinicians using this information might perhaps note that a more accurate label for this factor might be Perceived Social Support.

Results from this study have potential usefulness for clinicians in case conceptualization. In the process of naming the clusters, effort was taken in making a systematic and psychometrically congruent clinical analysis and description of the cluster profiles. The phrases used to describe the profiles were designed to be descriptive and avoid derogatory connotations. For instance, Interpersonally Focused reflects the dominance of the (perceived) Social Support scales without placing value. Similarly, the Stoic profile (one who suffers without assistance or complaint) acknowledges the high level of Punishing Responses in Factor 1/Suffering and the intermediate levels of Factor 3/Activity combined with low levels of Factor 3/Social Support. A similar rationale was used for describing cluster 3 as Adaptively Focused. The labels we use as clinicians to describe patients have an impact on our case conceptualization, our approach, and the understanding and approach of our staff. These labels are meant to be informative as well as descriptive to assist in guiding an understanding of the patient.

The large heterogeneous sample and use of mean scaled scores are strengths of this investigation. After applying minimal inclusion criteria, the sample included participants ranging in age from 13 to 92 years, ranging in duration of pain from
1 month to 50 years, and represented all IASP sites. The demographic and clinical diversity of the participants provides evidence for considerable generalizability and external validity.

Limitations of this study lie in the exclusion of participants representing certain populations. Due to the nature of the treatment facilities, participants with no insurance or ability to pay were excluded. Also excluded were participants who did not seek treatment in a large facility designed to treat chronic pain. Due to the nature of the treatment facilities, participants with no insurance or ability to pay were excluded. Also excluded were participants who did not seek treatment in a large facility designed to treat chronic pain. There are a significant number of patients with chronic pain who seek treatment solely from their family physicians. Additionally, most participants lived in the Chicago metropolitan area and identified themselves as Caucasian, representing a relatively truncated sample of regional and cultural influence. The investigation did not investigate the role of race, ethnicity, or sex. Similar to combining heterogeneous taxonomic groups, combining heterogeneous demographic groups without accounting for unidentified group differences can confound results. These factors may limit the application of the results to different regional and demographic groups.

The study used archival data, and was thus limited by the relatively few measures available for validation of the clusters. Ideally, evidence for validity of the clusters would benefit from significant differences among clusters on measures of suffering, social support, and activity.

Future investigations might include replicating these results and ascertaining the external and predictive validity of the revised MPI scales and the revised taxonomic system as well as its test–retest reliability.

Conclusion

This study is most accurately viewed as a source of information and/or a step in the continued progression toward a more definitive taxonomic system for chronic pain patients that will serve as a useful tool for case conceptualization, treatment planning, and predicting treatment outcomes. The suggested revisions were developed with a more representative sample than the old system, using more stringent methodology, and based on a more solid factor analysis of the instrument. Clinical utility, accessibility, and applicability were also incorporated as factors throughout. The suggested improvements could be easily programmed into another computer-based scoring system and perhaps an administration system as well allowing for a more convenient method for translating the revised scales and the discriminant function formulas into useful categorical results, similar to the current MAP system. If future research indicated the need, perhaps the MPI/MAP system could be revised to accommodate these changes.

References

13 Rudy TE, Turk DC, Zaki H, Curtain H. An empirical taxometric alternative to traditional classifica-
A Taxonomy of Chronic Pain Patients


