THE McGILL PAIN QUESTIONNAIRE
(Ronald Melzack, 1975)

Purpose

The McGill Pain Questionnaire (MPQ) was designed to provide a quantitative profile of three aspects of pain (1). The method was originally used in evaluating pain therapies, but it has also been used as a diagnostic aid (2).

Conceptual Basis

Melzack's major contribution to pain measurement was to emphasize that the pain experience comprises several distinct aspects. In 1973 he wrote:

The problem of pain, since the beginning of this century, has been dominated by the concept that pain is purely a sensory experience. Yet it has a unique, distinctively unpleasant, affective quality that differentiates it from sensory experiences such as sight, hearing or touch. . . .

The motivational-affective dimension of pain is brought clearly into focus by clinical studies on frontal lobotomy . . . Typically, these patients report after the operation that they still have pain but it does not bother them. . . . the suffering, the anguish are gone. . . . Similarly, patients exhibiting "pain asymbolia" . . . after lesions of portions of the parietal lobe or the frontal cortex are able to appreciate the spatial and temporal properties of noxious stimuli (for example, they recognize pin pricks as sharp) but fail to withdraw or complain about them. . . .

These considerations suggest that there are three major psychological dimensions of pain: sensory-discriminative, motivational-affective, and cognitive-evaluative. (3), pp93-95)

Melzack argued that these three aspects of the pain experience are subserved by distinct systems in the brain (3) and he sought to measure these dimensions with the MPQ. This link between emotions and neurological structures became a central feature of his neuromatrix model of pain, developed long after the MPQ, and which builds on his gate control theory (see the introduction to this chapter) (4).

Melzack stressed that the questionnaire represented a first attempt at developing a measurement reflecting his theory of pain, and he suggested that other investigators might ultimately refine it (1). Nonetheless, the method continues to be used in its original form.

Description

The complete MPQ comprises sections recording the patient's diagnosis, drug regimen, pain history, present pain pattern, accompanying symptoms and modifying factors, the effects of pain, and the list of words describing pain which is the part of the instrument most commonly used (5), Table 1). The present discussion concerns only this latter section of the overall questionnaire.

Melzack and Torgerson selected 102 words describing pain from the literature and from existing questionnaires (6). They sorted these words into the three major classes proposed in Melzack's theory of pain: words concerned with the sensory qualities of pain (e.g., temporal, thermal), those covering affective qualities of pain (e.g., fear, tension), and "evaluative words that describe the subjective overall intensity of the total experience of pain" (6). Within the three major classes, Melzack and Torgerson grouped words that were qualitatively similar, as shown in Exhibit 9.3. The suitability of this a priori grouping was checked by 20 reviewers. At first there were 16 such subclasses, but four others were added to give the final 20 subclasses (1).

An equal-appearing interval scaling procedure was used to estimate the intensity of pain represented by the words in each subclass. Three groups of judges (140 students, 20 physicians, and 20 patients) rated each word on a seven-point scale (6). Where there was disagreement among the three groups of judges on the rank ordering of a word within a subclass, the word was deleted from the questionnaire; scale values for the remaining words were based on the ratings made by patients. These values are shown in Exhibit 9.4.

The words were originally read to the patient by an interviewer so that unfamiliar words could be explained. This took 15 to 20 minutes, reducing to 5 to 10 minutes for patients familiar with the method. Hospitalized patients may require longer: a mean of 24 minutes has been reported (5), p36). Subsequent users have employed the checklist in a written format. The respondent is asked to select the one word in each subclass that most accurately describes her pain at that time (1). If none of the words applies, none is chosen. Other instructions have been used: patients may be asked to describe their "average pain," their "most intense" pain (7), or how their pain "typically feels" (8).

Four scoring methods were proposed by Melzack:

1. The sum of the scale values for all the words chosen in a given category (sensory, etc.) or across all categories. This was called a Pain Rating Intensity Score using scale values: PRI(S). The scale weights are shown in Exhibit 9.4; note that no scale weights are available for category 19. Separate PRI scores can be calculated for the sensory scales (range 0 to 42), for the affective scales (range 0 to 14), for the evaluative scales (range 0 to 5), and for the miscellaneous scales (range 0 to 17) (5), p37). The category scores and the PRI(S) can also be placed on a common scale by dividing the summed value by the maximum possible score, giving a range from 0 to 1.00 (9), p45).

2. As in 1, but replacing the scale values by a code indicating the rank placement of each word selected within its subclass: PRI(R). This score can be modified by multiplying the rank placement by a weight for each of the 20 groups of words in order to reflect the differing severity represented by the groups. The 20 weights were shown by Melzack et al. (10), Table 1).

3. The Number of Words Chosen: NWC.
4. The Present Pain Intensity (PPI) is a simple numerical 0-to-5 scale as shown in Exhibit 9.3 (although a 1-to-5 scale has also been used).

Reliability

Melzack reported a small test-retest study in which ten patients completed the questionnaire three times at intervals ranging from three to seven days; there was an average consistency of response of 70.3% (1), p287). The extent of correlation between sensory and affective scores varies according to the type of pain being considered: in one study the range was from 0.51 to 0.85 (11), Tables 2, 3).

Validity

Melzack reviewed the agreement among the four scoring methods (1). Correlations are presented in Table 9.2 and show that the PPI score does not agree closely with the other scores.

Dubuisson and Melzack compared responses to the MPQ given by 95 patients suffering from various pain syndromes. Discriminant function analyses showed that 77% of patients could be correctly classified into diagnostic groups on the basis of their verbal description of pain (2; 12).

Correlations between MPQ scores and visual analogue scales for 40 patients ranged from 0.50 for the affective score to 0.65 for the PPI and for the total score (13), Table 2). Equivalent figures from another study were 0.42 for the affective score and 0.57 for the evaluative (14), Table 1).

Several reviewers have addressed the question of whether Melzack's selection and grouping of words do indeed reflect the three dimensions he proposed. Studies in Canada (7), the United States (15), and Britain (16), each using different types of pain patient, have reviewed the factor structure of the MPQ. Critical reviews of factor analytic investigations were given by Prieto and Geisinger (17) and by Holroyd et al. (18). Analyses have extracted as few as three factors, and as many as seven. For example, two studies extracted four factors (13; 15), two found five (7; 19), and one found six (20). Holroyd et al. discussed methodological reasons for the discrepancies and then proceeded to analyze
data from a large sample of low back pain patients. They concluded that the sensory category of adjectives fell into two groups, separate from the affective and evaluative scales, making a four-factor solution. However, a single, higher-order factor accounted for two-thirds of the variance, suggesting that forming separate factor scores would have little utility (18), Table III and p307). Not surprisingly, differing conclusions have been drawn concerning the structure of the MPQ. Reading concluded that "Factorial investigations of the questionnaire provide support for the distinction between affective and sensory dimensions, but not for a distinctive evaluative component" (16). Crockett et al. concluded that "the repeated demonstration that the MPQ assesses substantially more than the three components originally proposed suggests that considerable caution is warranted with respect to using the 'sensory-evaluative-affective' method of scoring this test" (19), p461). Turk et al., however, concluded that the factor analyses confirmed the three-factor structure (sensory, affective, evaluative), although these factors were intercorrelated (21). While there is room for debate over the structure of the MPQ, the factors tend to intercorrelate highly (0.64 to 0.81), although Melzack argued that this does not contradict their use as separate scores (22). Torgerson applied multidimensional scaling to a subset of 17 descriptors and basically confirmed their placement in the groups shown in the Exhibit (23), Table 1).

The PPI scale correlated -0.68 with the pain questions from the World Health Organization’s quality of life scale, the WHOQOL. Regression analyses showed that it was principally the evaluative component of the MPQ that contributed to predicting quality of life (24), p401).

**Table 9.2 about here**

*Text same as Table 8.2, page 349, in 2nd edition*

### Alternative Forms

Wilkie et al. reviewed several variants on the overall McGill instrument that may cause confusion: the McGill-Melzack Pain Questionnaire, the McGill Pain Assessment Questionnaire and the McGill Pain Questionnaire, all of which differ, although all contain the pain descriptors reviewed here (5), Table 1).

A short-form MPQ has been developed which includes 15 words (11 sensory, four affective) (25). These words are identified with an asterisk in Exhibit 9.4; note that Melzack presented hot and burning together, as "hot-burning," and likewise combined "tiring-exhausting" and "punishing-cruel." Each word or phrase is rated on a four-point intensity scale: 0 = none, 1 = mild, 2 = moderate and 3 = severe. The six-point PPI rating is also included, as is a visual analogue scale (25), Figure 1). The short form was found to perform almost as well as the full MPQ in a study of cancer pain (26). Other abbreviations of the MPQ can be found on the web; it is frequently not clear whether these have been validated.
Because of the linguistic nuances in the MPQ, translation is challenging. Nonetheless, there are versions in most European languages including French (27), Italian (28), German (29), Dutch (30; 31), Danish (32), Norwegian (33), and Finnish (34). There is an Arabic version (35) and a Mexican Spanish version has been tested (36). Many of these adapt, rather than translate, the original and include widely varying numbers of adjectives: results cannot be compared across them. A Swedish abbreviated version, for example, contains 15 items and appears to have adequate reliability and validity (37). At times the linguistic challenge of translating the MPQ seems too strong to resist: there are at least three French Canadian versions, as well as a French-French version. To the surprise of few, a study from Paris declared the version from France to be superior (27).

There have been various extensions to the MPQ. One version, called the McGill Comprehensive Pain Questionnaire, covers details of the patient’s illness, personality, milieu, and coping resources (38). The Dartmouth Pain Questionnaire adds questions on pain complaints, somatic pain symptoms, and reductions in function and in self-esteem; some validity evidence is available (39). Another version added computer animation to illustrate categories such as pressure (picture of a clamp squeezing a ball); throbbing (a hammer falling from various heights); piercing (a needle), and burning (40).

Reference Standards

Wilkie et al. analyzed pain scores from studies of patients with different types of pain: cancer, chronic back pain, postoperative, labor, dental, etc. They presented mean pooled scores for seven pain conditions as follows: PRI-total = 23; PRI-sensory = 13.9; PRI-affective = 3.3; PRI-evaluative = 2.5 and PRI-miscellaneous = 4.9; NWC = 9.2 and PPI = 2.3 (5), Table 2). They then presented profiles for each condition. For example, low back pain patients had a high mean PRI-sensory score of 16.3, compared to cancer patients (12.1) or dental patients (10.7), but overall, the differences between conditions were small (5), p39). Wilkie et al. also reported the frequency of reporting particular adjectives by each type of pain (5), Table 3). There were no clear patterns in the words chosen by patients with particular types of pain.

Commentary

Melzack's McGill Pain Questionnaire is the leading instrument for describing the diverse dimensions of pain. It was based on an explicit theory of pain. It is used extensively in many countries and has taken on the status of gold standard against which newer instruments are compared. Sections of the MPQ have been incorporated into other scales, such as the Brief Pain Inventory, the MOS Pain Measures, the Back pain Classification Scale, and the Pain Perception Profile reviewed in this chapter. Nonetheless, Melzack originally considered the MPQ as a preliminary version of a measurement method rather than a definitive scale, and variants have been used in several studies. A limitation of using verbal descriptors is that people’s use of pain language is probably not sufficiently consistent for the MPQ to be used to diagnose types of pain. In analyses of groups of patients, responses for

different types of pain are distinguishable, but it is unlikely that an individual's responses could be used to diagnose the cause of pain (20). This is especially true in patients with psychological disturbance (41).

Discussions of the MPQ have raised four main issues: the constitution of subcategories within the questionnaire; the suitability of selecting only one word from each category; the choice of summary scores, and the question of whether the MPQ does, in fact, reflect Melzack's theory of pain. 

The approach of selecting only one word in each category assumes that the words in each category are homogeneous and differ only in intensity. Some of the early studies by Reading et al. suggest that this approach is suitable (42), p381), although this would appear to be a suitable topic for IRT. In terms of scoring the MPQ, the 20 category scores proved superior to the three subscale scores in one study (42). Hand and Reading found no advantage in an alternative scoring approach that used 16 instead of 20 categories (43).

We may enquire whether the MPQ does, in fact, succeed in reflecting Melzack's theory of pain. Crockett et al. argued that the results of their factor analyses offered empirical validation of Melzack's a priori classification of pain descriptors and concurred with Melzack's emphasis on the need to describe pain in terms of several dimensions rather than as a single intensity score. However, there remains debate over the structure of the MPQ, and technical problems in examining how closely it reflects Melzack's theory have enraptured methodologists. For example, there are several problems in using factor analysis to assess the validity of the questionnaire. Evidently the conventional response procedure cannot be used whereby the respondent selects only one word in each subclass, for then the correlations among the words will be zero, and so the grouping of words into subclasses cannot be tested empirically. Melzack and Torgerson recognized that words from different components (e.g., affective, evaluative) may correlate with one another, while different words in each group, or different subclasses in each component, will not necessarily intercorrelate (6). If this is the case, one would not expect words in, say, the sensory component to load on a single factor. Furthermore, because each word reflects both a type and an intensity of pain, a factor analysis may extract type of pain or intensity of pain factors, or both. This was illustrated by a study that departed from normal usage and presented the MPQ words in random order, asking subjects to check every word that described their pain (8). The seven factors that were interpretable cut across Melzack's groupings and included words at similar levels of intensity from a wide range of subclasses.

Although it is hard to show that Melzack's questionnaire does reflect his conceptual definition of pain, the MPQ is still the leading pain measurement scale, and we recommend its continued use. On a lighter note, readers should always be wary of the information contained in apparently authoritative web sites. One (accessed in the fall of 2004) that may best remain nameless attributes the MPQ to an author named McGill who could be located in the United Kingdom (to judge from the name, perhaps Scotland?) Lest anyone else may be confused, Ronald Melzack is an emeritus professor of psychology at McGill University in Montreal, Canada. Aside from his work on pain, he is also the author of books on Eskimo folklore.
References


