

Assessment of emotional functioning in pain treatment outcome research

Robert D. Kerns, Ph.D.

VA Connecticut Healthcare System

Yale University

Running head: Emotional functioning

Correspondence: Robert D. Kerns, Ph.D., Psychology Service (116B), VA Connecticut Healthcare System, 950 Campbell Avenue, West Haven, CT 06516; Phone: 203-937-3841; Fax: 203-937-4951; Electronic mail: robert.kerns@med.va.gov

Assessment of Emotional Functioning in Pain Treatment Outcome Research

The measurement of emotional functioning as an important outcome in empirical examinations of pain treatment efficacy and effectiveness has not yet been generally adopted in the field. This observation is puzzling given the large and ever expanding empirical literature on the relationship between the experience of pain and negative mood, symptoms of affective distress, and frank psychiatric disorder. For example, Turk (1996), despite noting the high prevalence of psychiatric disorder, particularly depression, among patients referred to multidisciplinary pain clinics, failed to list the assessment of mood or symptoms of affective distress as one of the commonly cited criteria for evaluating pain outcomes from these programs. In a more recent review, Turk (2002) also failed to identify emotional distress as a key index of clinical effectiveness of chronic pain treatment. A casual review of the published outcome research in the past several years fails to identify the inclusion of measures of emotional distress in most studies of pain treatment outcome, other than those designed to evaluate the efficacy of psychological interventions. In a recent edited volume, The Handbook of Pain Assessment (Turk & Melzack, 2001), several contributors specifically encouraged inclusion of measures of psychological distress in the assessment of pain treatment effects (Bradley & McKendree-Smith, 2001; Dworkin, Nagasako, Hetzel, & Farrar, 2001; Okifuji & Turk, 2001).

A primary goal of this monograph is to further encourage the routine inclusion of specific measures of emotional functioning in the design of pain treatment outcome studies, including those designed to evaluate the efficacy of pharmacological

interventions for pain management, and to provide specific recommendations for the selection of appropriate measures of emotional functioning in such studies. This monograph will begin with the presentation of a rationale for the importance of assessing emotional functioning in this context. Specific dimensions of the broader experience of affective distress will be highlighted, particularly the experiences of depression, anxiety, and anger. The importance of discriminating mood states, from mood symptom clusters, from psychiatric disorders will be discussed. This discussion will be followed by a comprehensive review of the key measures of these constructs and the data related to their reliability, validity, and utility. Recommendations will be offered for the selection of specific measures of emotional functioning in pain outcome research. The paper will conclude with a few specific suggestions for future research in this area.

Why assess emotional functioning in pain treatment outcome research?

Kerns, as recently as 1996, at a gathering of experts to discuss future directions for pain management, argued that psychosocial variables, including measures of emotional functioning, should be considered to be of primary importance, rather than continuing to be viewed as secondary to measures of pain relief, per se. In support of this perspective, Kerns argued that the dominant contemporary models of pain emphasize the multidimensional nature of the experience of pain (Flor, Birbaumer, & Turk, 1990; Turk, Meichenbaum, & Genest, 1983) and that the sole reliance on pain reports is inadequate for capturing the breadth and complexity of the experience of pain. He also cited the high prevalence and enormous costs associated with the experience of emotional distress and disorder among persons with persistent pain (e.g., Atkinson, Slater, Patterson, Grant,

& Garfin, 1991; Banks & Kerns, 1996; Dworkin & Caligor, 1988; Gaskin, Greene, Robinson, & Geisser, 1992). Finally, support comes from empirical demonstrations of the relative independence of pain and emotional distress (DeGagne, Mikail, & D'Eon, 1995; Holroyd, Malinoski, Davis, & Lipchik, 1999; Linton & Goettestam, 1985; Mikail, DuBreuil, & D'Eon, 1993) and from studies that have identified subgroups of persons with chronic pain on the basis of their high level of reported interpersonal and emotional distress (Jamison, Rudy, Penzein, & Mosley, 1994; Klapow, Slater, Patterson, Atkinson, Weickgenant, Grant, & Garfin, 1995; Klapow, Slater, Patterson, Doctor, Atkinson, & Garfin, 1993; Turk & Rudy, 1988; 1990).

Dworkin and his colleagues (2001) have also asserted that it may be particularly important to assess emotional functioning in clinical trials to control for the possible effects of the intervention on pain via an intermediate effect on psychological distress, including sleep dysfunction or anergia. These authors suggest that this may be particularly important when evaluating the effects of pharmacological interventions that are known to have psychological or behavioral effects, and they cite research investigating the analgesic effectiveness of tricyclic antidepressants as an example. Similar concerns may be raised in the context of nonpharmacological trials, including for example, studies of the efficacy of exercise or cognitive-behavior therapy that are likely to have positive effects on mood and other symptoms of affective distress.

Dimensions of emotional functioning

An extensive empirical literature from both laboratory and clinical settings highlights the important relationship between mood states and symptoms of emotional distress and the

experience of pain (cf., Robinson & Riley, 1999). The three most commonly studied dimensions of negative emotion are anxiety, depression, and anger. Studies generally have focused on developing estimates of the prevalence of mood disorders among persons with chronic pain, the development and refinement of models that describe and explain the impact of negative emotions on pain and pain-related disability and vice versa, and strategies for reliable assessment of these constructs.

The experience of anxiety and fear in association with the experience of pain is almost ubiquitous. An anxious mood state has long been recognized as having a dramatic and reliable effect on pain perception in the laboratory setting (Chaves & Barber, 1976; Mersky, 1978). Rates of anxiety disorders have been consistently found to be high among sample of persons with chronic pain (Asmundson, Jacobson, Allerdings, & Norton, 1996; Atkinson et al., 1991; Fishbain, Goldberg, Maegher, & Steele, 1986). There is also evidence that specific anxiety disorders may be particularly common among persons with certain painful medical conditions. For example, although rates of panic disorder in heterogeneous samples of persons with chronic pain have been reported to be in the range of 16% (Katon, Egan & Miller, 1985), Beitman and colleagues have reported that between 34% and 59% of persons with chest pain of unknown etiology may meet diagnostic criteria for panic disorder (Beitman, Mukerji, Kushner, Thomas, Russell, & Logue, 1991).

Beckham and colleagues (1997) have reported that up to 80% of Vietnam veterans with post-traumatic stress disorder (PTSD), another specific anxiety disorder, report chronic pain (Beckham, Crawford, Feldman, Kirby, Hertzberg, Davidson, & Moore, 1997), and other studies have documented rates of PTSD to be as high as 50 to

100% among persons receiving treatment at pain treatment centers (Kulich, Mencher, Bertrand, & Maciewicz, 2000). Recently, presence of persistent, diffuse musculoskeletal pain has been identified as one of the cardinal features of Gulf War Syndrome, a multisymptom complex associated with significant physical and mental health dysfunction (Donta, Clauw, Engel et al., 2003), and increased physical symptoms, including muscle aches and back pain, have been shown to be more common among Gulf War veterans with PTSD relative to those without PTSD (Baker, Mendenhall, Simbartl, Magan, & Steinberg, 1997).

Of growing interest are observations of the specifically important role of pain-relevant anxiety, and in particular, a specific phobia, that is fear of pain (Fernandez, 2002). Patterns of pain-related fear and behavioral avoidance related to pain have been observed to effect reports of pain and pain-related disability as well as physical performance measures (Asmundson, Jacobson, & Allardings, 1997; Lethem, Slade, Troup, & Bentley, 1983; Linton et al., 1985; McCracken, Gross, Aikens, & Carnrike, 1996; McCracken, Gross, Sorg, & Edmands, 1993; Waddell, Newton, Henderson, Somerville, & Main, 1993).

A particularly high rate of the coprevalence of pain and depression is well documented (Banks & Kerns, 1996; Romano & Turner, 1985), as is evidence that depression among persons with chronic pain may be associated with increased healthcare system utilization and increased disability. Romano and Turner (1985) noted that reported prevalence rates of depression range from as low as 10% to as high as 100%, and Banks and Kerns (1996) have suggested that, on average, rates of depression among persons presenting for multidisciplinary pain treatment are approximately 50%. These

later authors have suggested that the rates of depressive disorder are higher among persons with chronic pain than among persons with any other acute or chronic illness. There is also evidence that the presence of depression may negatively influence response to treatment (Haythornthwaite, Seiber & Kerns, 1991; Kerns & Haythornthwaite, 1988). These observations have led to an extensive body of research designed to examine putative neurobiological and psychosocial mediators of the relationship between pain and depression.

Anger among persons with pain has also been widely observed (Fernandez, 2002), has been found to be a particularly strong correlate of pain intensity, even relative to other negative emotional states (Summers, Rapoff, Varghese, Porter, & Palmer, 1991), and has also been demonstrated to interfere with treatment (Fernandez & Turk, 1995). The style of expressing intense negative emotion, particularly anger, has been hypothesized to play a role in the development and perpetuation of persistent pain, pain-related disability, and depression (Beutler, Engle, Oro-Beutler, & Daldrup, 1986; Catchlove & Braha, 1985; Kerns, Rosenberg, & Jacob, 1994). Perhaps due to the relative lack of attention to anger in psychiatric nomenclature, research on the prevalence of problems with anger among chronic pain samples, and empirical research designed to investigate the relationship between anger and pain remains in their relative infancy.

Conceptual and empirical challenges in the assessment of anxiety, depression, and anger

Emotions are subjectively experienced private events that vary in intensity and are generally experienced as positive, or pleasant, or negative or unpleasant. There continues

to be no agreement among theoreticians, scholars, and empiricists about the number of emotions or an acceptable taxonomy for their characterization or labeling. Fernandez (2002) listed several specific emotions as being of particular clinical interest: joy, anger, fear, sadness, shame, guilt, and envy. Among these, anxiety, depression, and anger have drawn the greatest attention among clinicians and researchers in the field of pain.

Fernandez (2002) also emphasizes the differences between emotions as a discrete episode, and mood as a relatively continuous process. These phenomena are further distinguished from temperament and personality, that is the tendency to experience certain emotions at a relatively high frequency, or certain moods for extended periods of time. The term “trait” is often used to describe this tendency, whereas “state” commonly refers to the momentary feelings. Among measures of anxiety, the most commonly used measure, the State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Luschene, 1970), is one that includes both a “state” form and a “trait” form that attempt to discriminate between transient experiences of anxiety and a more general tendency to be anxious. As will be discussed in more detail below, it is important to distinguish the experience of negative mood from a pathological state or disorder of emotion or mood. An affective disorder, for example, is characterized by an intensity and/or frequency that leads to an experience of dysfunction or concern on the part of the person experiencing the disorder or significant others.

The nature of the relationship between pain and emotion has been the target of considerable attention. Fernandez (2002) provides an overview of the several hypothesized models that can be articulated to describe, if not explain, the nature of the relationship. Each model generally attempts to ascribe a temporal relationship between

pain and affect, and some incorporate a notion of causal direction. Perhaps the model with the greatest support is the simplest or most parsimonious one, as well. This model suggests simply that pain and emotion are correlates. Five other dynamic models of the relationship between pain and affect hypothesize that (1) affect is a predisposing factor in the experience of pain, (2) affect is a precipitating factor in pain, (3) affect is an exacerbating factor in pain, (4) affect is a consequence of pain, and (5) affect as a perpetuating factor in pain.

The ability to discriminate among these negative emotional experiences has been the target of extensive discussion and empirical investigation. In the study of clinical pain phenomenon, in virtually all studies that have included measures of more than one of these constructs a high level of intercorrelation among the measures has been demonstrated (e.g., Gaskin, et al., 1992; Summers et al., 1991; Wade, Price, Hamer, & Schwartz, 1990). Brown and colleagues (1996) employed structural equation modeling to control for measurement error and demonstrated that measures of anxiety, depression, and anger formed a single affective distress factor. Gracely (1992) has similarly asserted that negative emotions are conceptually and empirically interrelated. Robinson & Riley (1999) have argued that the co-occurrence of these negative emotional states among persons with chronic pain, in conjunction with empirical observations of the high degree of intercorrelations among measures of these emotions, represents a serious challenge to those interested in making discriminations among them. Attention to issues of method variance (i.e., the measurement of constructs using similar methods, for example, self-report methods) and response set or bias in reporting has been highlighted as important in attempts to improve the discrimination among measures of negative emotions in the

context of the assessment of the chronic pain experience (Holroyd, Talbot, Holm, Pingel, Lake, & Saper, 1996; Robinson, Myers, Sadler, Riley, Kvaal, & Geisser, 1997).

The importance of discriminating among the transient experience of a negative mood state, the experience of a cluster of symptoms commonly associated with emotional distress, and the diagnosis of a psychiatric disorder is critical in a discussion of the assessment of emotional functioning among persons with pain. The construct of “depression”, its operationalization, and its measurement represent the most common example of this challenge. The experience of sadness or even frank depression may be reliably and validly measured by a self-report questionnaire that asks persons to endorse a set of adjectives commonly accepted as representing this mood state (e.g., the Profile of Mood States; McNair, Lorr, & Doppelman, 1971). In contrast, the two most commonly employed self-report measures of the broader construct of depressive symptom severity are the Beck Depression Inventory (BDI; Beck, Ward, Mendelsohn, Mock, & Erbaugh, 1961) and the Center for Epidemiological Studies – Depression scale (CES-D; Radloff, 1977) that are comprised of numerous items reflecting not only the state of depressed mood, but commonly associated symptoms such as sleep difficulties, loss of interest in pleasurable activities, and loss of appetite and weight.

The diagnosis of Major Depressive Disorder (MDD), a particularly common psychiatric diagnosis among persons with chronic pain, requires the use of a clinical interview to assure that several specific criteria for the diagnosis are met. These include: (1) the presence of depressed mood or loss of interest or pleasure; (2) the presence of at least three additional symptoms such as insomnia, fatigue or loss of energy, or recurrent thoughts of death from a list of nine symptoms; (3) the presence of these symptoms for at

least a two week period of time and that represents a change from previous functioning; (4) the symptoms cause clinically significant distress or impairment in social role functioning; (5) the symptoms are not due to the direct physiological effects of a substance or a general medical condition; and (6) the symptoms are not caused by bereavement. Structured psychiatric interviews and diagnostic decision trees have been developed for facilitating reliable and valid diagnosis. As described below, the sensitivity and specificity of symptom severity measures such as the BDI and CES-D in making a diagnosis of MDD have also been described.

Sullivan (2001) has argued that diagnosis of psychiatric disorders should not be considered only after a “medical disorder” has been “ruled out” as the cause for persistent pain. He suggests that such a process leads to unnecessary and costly medical diagnostic procedures, iatrogenic injury, and poor clinical management. As an alternative, Sullivan suggests that psychiatric disorder should be considered any time a pain disorder becomes chronic. In support of this argument, Sullivan and others cite an extensive epidemiological literature that documents a high prevalence of primary psychiatric disorder among persons with chronic pain. Most commonly cited are depressive disorders, anxiety disorders, particularly panic disorders and post-traumatic stress disorder (PTSD), substance abuse and dependence disorders, somatization disorder, and personality disorders. An extensive body of research has attempted to address the question of whether pain most commonly leads to the development of subsequent psychiatric disorder (e.g., Brown, 1990), versus whether psychiatric disorder serves as a vulnerability or predisposing factor for the development of persistent pain conditions (e.g., Katon, Egan, & Miller, 1985). Ultimately, Von Korff & Simon (1996) have

proposed that pain and psychiatric disorder should be viewed as reciprocal processes of illness expression and social adaptation.

It is important to acknowledge the complexity of the discriminations that may be necessary in making a determination about the presence or absence of frank psychiatric disorder. For example, it is clear that many persons with chronic pain may experience significant alterations in mood and even develop additional symptoms of depressive disorder without meeting the diagnostic criteria for a psychiatric disorder. Alternatively, persons with chronic pain may also develop anhedonia, or a pervasive loss of interest in previously enjoyed activities, without experiencing depressed mood. If the other criteria of MDD are met, these individuals may be appropriately diagnosed with this disorder. Distinguishing MDD from other psychiatric disorders such as Dysthymic Disorder, anxiety disorders, and adjustment disorders represents an additional challenge for the clinician and investigator.

Another important question that has led to extensive discussion and some empirical research is whether a model of depressive disorder that is inclusive of “physical symptoms” that may be attributable to the experience of pain (e.g., insomnia, fatigue) versus a diagnostic conceptualization that excludes these symptoms is more appropriate and valid (Geisser, Roth & Robinson, 1997; Rodin & Voshart, 1986). The question has been raised in the context of reviews of epidemiological studies of the prevalence of MDD where some have suggested that rates of psychiatric disorder have been inflated and by others who suggest that treatment for the psychiatric disorder may be inappropriate and ineffective unless or until the pain condition is addressed. Koenig, George, Peterson & Pieper (1997) have argued that the evidence supports the reliability

and sensitivity or “inclusive” models of MDD, in particular, and provide a compelling review suggesting that somatic symptoms of this disorder are not a direct function of the experience of pain.

Measures of emotional distress

The following review of measures of emotional distress begins with a consideration of primarily self-report measures of anxiety, particularly measures of pain-related fear, depressive symptom severity, and anger that may have relevance for the evaluation of interventions for pain. Not reviewed are single item measures of emotional functioning (e.g., visual analogue and numeric rating scale measures) since these strategies for the assessment of emotional distress, although appearing with some frequency in the pain literature, have largely been displaced by other multi-item standardized measures that have substantial evidence of reliability and validity. Also not considered are measures of other domains of emotional distress, including emotions other than anxiety, depression, and anger, since these have not generally been the target of particular interest and investigation in the pain field. Notable in their absence, however, are measures of marital and family distress that represent an increasingly important and interesting area of investigation in the pain literature. The decision not to review these measures is because relational distress has not become a primary target of intervention in the pain field, and when this dimension of distress has been investigated, it has largely not been affected by pain treatment. The review of measures of anxiety, depression and anger is followed by a review of several multidimensional measures of emotional functioning and distress. The section concludes with a brief review of two semistructured

psychiatric diagnostic interviews that serve as the primary methods for reliable diagnosis of disorders of emotional functioning.

Unfortunately, with the exception of measures of pain-related anxiety and fear that have recently been developed, virtually none of the other measures were developed for use in the assessment of emotional distress among persons with clinical pain conditions. Furthermore, most were developed with the intent of characterizing or quantifying the presence and severity of emotional distress or for use in screening for the presence of psychiatric disorder. Nevertheless, in several cases, these measures have subsequently been evaluated for their roles and utility as measures of emotional distress among persons with pain, and their value as measures of change in levels of emotional distress has been investigated in the context of studies of the efficacy of pain interventions.

Measurement of anxiety

The measurement of anxiety in the field of pain and pain management is increasingly dominated by measures of the construct of fear of pain. Recent data suggest that pain-related fear may be a key component in the development and maintenance of pain-related physical disability (McCracken, Faber, & Janeck, 1998). Pain-related fear (also referred to as *kinesiophobia*) may be defined as the constellation of fearful feelings and avoidance behaviors in anticipation of a re-experiencing of painful sensations or of a re-injury (Kori, Miller, & Todd, 1990). Research has demonstrated that for some individuals with chronic pain pain-related fear may mediate treatment-related improvement (Vlaeyen, de Jong, Geilen, Heuts, & van Breukelen, 2001). A brief

consideration of more general measures of anxiety will be considered, followed by more detailed reviews of three measures of fear of pain, namely the Pain Anxiety Symptoms Scale (PASS; McCracken, Zayfert, & Gross, 1992), the Tampa Scale of Kinesiophobia (TSK; Kori, Miller, & Todd, 1990), and the Fear-Avoidance Beliefs Questionnaire (FABQ; Waddell, et al., 1993). A fourth measure, the Fear of Pain Questionnaire – III (FPQ-III; McNeil & Rainwater, 1998), is not reviewed because of its recent development and the relative absence of empirical research that addresses its reliability and validity. Although none of these measures have been employed in an evaluation of pain treatment, they are briefly reviewed here because of their potential importance as pain-specific alternatives to the Spielberger measure of more general anxiety.

Spielberger State-Trait Anxiety Inventory (STAI)

By far the most commonly used measure of anxiety in the pain literature has been the Spielberger State-Trait Anxiety Inventory (STAI; Spielberger et al., 1970). The measure was specifically designed to aid in the discrimination of situational (state) anxiety and dispositional (trait) anxiety. The STAI consists of two 20-item self-report inventories of each of these constructs. Respondents rate the degree of agreement with brief statements (e.g., “I feel calm”) on 4 point scales ranging from “not at all” to “very much so” in terms of either their present state or their frequency over time (trait version). There is a high concordance between pain and anxiety as measured by the STAI (Polatin, Kinney, Gatchel, Lillo, & Mayer, 1993), and it has been widely used as a pain outcomes measure. It has acceptable psychometric properties (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983), and it is sensitive to change (Mongini, Defilippi, & Negro, 1997).

Concerns about its sensitivity as an anxiety measure relative to its associations with other dimensions of emotional distress have been raised. The STAI has also been used as a measure of change in anxiety as a function of pain treatment and, in several studies, has been found to be sensitive to change (e.g., Applebaum, 1988; Bradley, Young, et al., 1987; Kerns, Turk, Holzman, & Rudy, 1986; Nielsen, Walker, & McCain, 1992; Richardson, Richardson, Williams, Featherstone, & Harding, 1994).

Pain Anxiety Symptom Scale (PASS)

The Pain Anxiety Symptom Scale (PASS; McCracken et al., 1992) was designed to assess the cognitive, physiological, and behavioral domains of pain-related fear. It includes 53 items distributed across four subscales measuring Fear of Pain, Cognitive Anxiety, Somatic Anxiety, and Escape and Avoidance. Respondents use 0 (never) to 6 (always) scales to endorse the frequency of each of the symptoms. The PASS has been demonstrated to have adequate internal consistency (McCracken et al., 1992) with indices of internal consistency ranging from 0.81 to 0.89 for each of the four scales, and 0.94 for the total scale. Good predictive validity (McCracken et al., 1998), and acceptable validity (McCracken, Gross, Aikens, & Carnrike, 1996) have also been demonstrated. The PASS has been criticized for its poor prediction of disability relative to other pain-related fear measures (Crombez, Vlaeyen, Heuts, & Lysens, 1999) and its factor structure has also been challenged ((Larsen, Taylor, & Asmundson, 1997).

The Tampa Scale of Kinesiophobia (TSK)

The Tampa Scale of Kinesiophobia (TSK; Kori et al., 1990) is a 17-item instrument with items assessing pain-related fear of movement or of pain sensations due to concerns about injury or reinjury (Kori et al., 1990). Recent data suggest that the TSK may be a better predictor of a range of pain symptoms and behaviors than the other pain-related anxiety scales (Crombez et al., 1999), and it has been found to be a better predictor of disability than pain intensity, biomedical signs and symptoms, or negative emotionality measures (Crombez et al., 1999; Vlaeyen et al., 1999).

Fear-Avoidance Beliefs Questionnaire (FABQ)

The FABQ (Waddell et al., 1993) is a 21 item self-report measure based on fear theory and avoidance behavior and was specifically designed to assess patients' beliefs about the effects of activity and work on the experience of pain. Five statements are included about the relationship between pain and physical activity and respondents use a 0 (completely disagree) to 6 (completely agree) scale to rate their endorsement of the statement. Eleven additional items reflect beliefs about the relationship between pain and work. Instructions require respondents to use a similar seven-point scale. The authors of the measure demonstrated two subscale scores related to these two domains, although other investigators have reported three distinct factors (Pfungsten, Kroner-Herwig, Leibing, Kronshage, & Hildebrandt, 2000; Pfungsten, Leibing, Harter, Kroner-Herwig, Hempel, Kronshage, & Hildebrandt, 2001). Results of each of these groups, as well as others (Buer & Linton, 2002; Crombez et al., 1998; Klenerman, Slade, & Stanley, 1995), generally support the validity of the measure as a predictor of behavioral performance and treatment outcome. Buer and Linton (2002) recently suggested that accumulating

evidence suggests that fear-avoidance beliefs may be an appropriate target for intervention. To date, the measure has not been used to evaluate outcome following pain treatment.

Measures of depressive symptom severity

The two most commonly employed measures of depressive symptom severity are the Beck Depression Inventory (BDI; Beck et al., 1961; Beck & Steer, 1993) and the Center for Epidemiological Studies – Depression scale (CES-D; Radloff, 1977). Both measures have strong evidence of reliability, stability, and validity for use among the general population and among persons with known psychiatric disorder, and both have been employed extensively in the pain literature, including use in studies of pain treatment outcome. Numerous additional measures of depressed mood and depressive symptom severity have also been developed, and several of these have been employed to a limited extent in the pain literature.

Beck Depression Inventory (BDI)

The BDI was developed to measure the behavioral manifestations of depression in adolescents and adults and to standardize the assessment of depressive symptom severity in order to monitor change over time (Beck et al., 1961). Items of the BDI were derived from observations of persons with depression made during psychoanalytic psychotherapy. In its original form, the BDI consisted of 21 groups of four to five statements describing symptoms in each cluster from low to high. Respondents were instructed to endorse the single item in each group that best describes how they were

feeling “right now”. The original version was designed to be used in an interview format, but subsequent versions have more commonly been used in a self-report, questionnaire format. An abbreviated version of the measure that included only 13 items was published in 1976. In 1978 the full scale was revised to eliminate redundancy among some of the items and the time frame for assessment was altered to “during the last week, including today”. Only four possible responses for each symptom cluster are now included, so that scores on the measure range from 0 to 63. In 1996, the BDI-II was published and included revisions to some items and the time frame for assessment to be consistent with the DSM-IV. Although the BDI-II has advantages in terms of the content of the items and consistency with current diagnostic nomenclature, concerns have been raised about the sensitivity to change during brief periods of time as a function of the lengthened time frame for assessment (Yonkers & Samson, 2000). The 21 item version of the BDI takes about 5 to 10 minutes to complete.

The reliability and several dimensions of validity of the measure have been extensively reported. In a review of 25 years of research with the BDI, Beck and colleagues reported on 25 studies that reported on the internal consistency of the measure (Beck, Steer, & Garbin, 1988). Across psychiatric, healthy, and medically ill samples, indices of internal consistency (alphas) ranged from 0.73 to 0.95. Stability estimates (i.e., test-retest correlations) have consistently been high as well, typically varying in the 0.80 to 0.90 range depending on the assessment interval and sample. Validity estimates for psychiatric patients have been assessed by examining the correlation between BDI scores and clinical ratings of depression (e.g., using the Hamilton Rating Scale for Depression) average about 0.72. For nonpsychiatric patients, the average validity estimate is 0.60.

Correlations with other common self-report measures of depressive symptom severity are reported to be 0.76 for the Symptom Checklist-90 and 0.60 with the Depression scale of the MMPI (Beck, Steer, & Brown, 1986; Yonkers & Samson, 2000). In a review of eight studies of sensitivity to change, Moran and Lambert (1983) found that the BDI was sensitive to change as a function of psychotherapy and pharmacotherapy outcome studies. Some evidence suggests a reporting bias for certain populations, including women, adolescents, and elderly persons, although the robustness of these observations is not clear.

The BDI has been used extensively in studies designed to evaluate the efficacy of pharmacologic and nonpharmacologic treatments for chronic pain (Applebaum, 1988; Burns, Johnson, Mahoney, Devine, & Pawl, 1998; Kerns et al., 1986; Khatami & Rush, 1982; Kleinke, 1992; Marhold, Linton, & Melin, 2001; Nicholas, Wilson, & Goyen, 1992; Phillips, 1987; Richardson et al., 1994; Tanzer, Melzack, & Jeans, 1986; Turner, 1982; Williams, Richardson, Nicholas, et al., 1996), and there is ample evidence of its sensitivity to change. Results of most studies provide compelling support for the use of the BDI in assessing improvements in depressive symptom severity as a function of pain treatment.

Center for Epidemiologic Studies – Depression Scale (CES-D)

The CES-D was developed to screen for the presence of depressive illness and to measure levels of symptoms of depression in community samples (Radloff, 1977). Items were selected from existing scales (e.g., BDI, MMPI Depression scale, Zung Self-Rating Depression Scale) to represent the major components of depression on the basis of

clinical and empirical studies. The measure includes 20 items that measure depressed mood, feelings of worthlessness, feelings of helplessness, loss of appetite, poor concentration, and sleep disturbance. Respondents are asked to rate the frequency of each symptom on a 0 (rarely or none of the time) to 3 (most or all of the time) scale with reference to the past week. Four items are worded in the positive direction to partially control for response bias. Scores on the measure range from 0 to 60. The CES-D takes about 5 minutes to complete (Radloff & Locke, 2000).

Indices of internal consistency (Cronbach's alpha) have been reported to be 0.85 for community samples and 0.90 in psychiatric samples. Split-half reliabilities are also high, ranging from 0.77 to 0.92. Test-retest correlations over a 6 to 8 week period range from 0.51 to 0.67 (Radloff & Locke, 2000). Roberts (1980) reported that studies of African-American and Mexican-American respondents revealed similar reliability estimates. The reliability and validity of the measure have also been examined in Asian-American, French, Greek, Hispanic, Japanese, and Yugoslavian populations (Naughton & Wicklund, 1993) and it has been translated into several other languages, as well (Radloff & Locke, 2000). Overall, high levels of internal consistency have been reported across numerous samples from the general population and patient samples, irrespective of age, gender, race, and geographic location. In a sample of chronic pain patients, the level of internal consistency was found to be 0.90 (Arnstein, Caudill, Mandle, Norris, & Beasley, 1999).

Indices of criterion-related validity have generally been reported to be moderate to high. For example, correlations between the CES-D and the Depression scale of the SCL-90 for samples of psychiatric patients have been reported to range from 0.73 to 0.90.

Correlations with the Hamilton Rating Scale for Depression for similar samples have ranged from 0.49 to 0.85. Studies of elderly samples have revealed somewhat lower validity estimates. Among samples of elderly persons, correlations between the CES-D and the Geriatric Depression Scale, and with the Zung Self-Rating Scale, have been reported to be only fair (DeForge, & Sobal, 1988; Gerety, Williams, Mulrow et al., 1994).

In a study of the CES-D in a primary care sample, the investigators provided evidence of the ability of the measure to discriminate between mild and severe, but not mild and moderate, or between moderate and severe, depression (Fechner-Bates, Coyne, & Schwenk, 1994). Sensitivity to change as a function of treatment for depression has been demonstrated (Weissman, Sholomskas, Portenger et al., 1977). On the other hand, this same group of investigators on the basis of another study concluded that the ability of the CES-D to discriminate among patients with “primary” depression from those with depression associated with other psychiatric disorders (e.g., substance abuse) was not optimal. Other investigators have similarly challenged the diagnostic sensitivity of the measure (Boyd, Weissman, Thompson et al., 1982; Breslau, 1985; Fechner-Bates, et al., 1994; Myers & Weissman, 1980; Roberts & Vernon, 1983). Investigators in the pain field have called for modifications of the measure in terms of item content (e.g., Blalock, DeVellis, Brown, & Wallston, 1989; Brown, 1990) or scale cut-offs for the diagnosis of depression (e.g., Magni, Caldieron, Rigatti-Luchini, & Merskey, 1990; Turk & Okifuji, 1994). Ultimately, it is fair to say that the measure lacks the sensitivity and specificity for supporting its use in clinical diagnosis without concurrent use of a psychiatric interview. The CES-D has increasingly been used for the assessment of outcome

following pain interventions, and in numerous cases, the measure has been demonstrated to be sensitive to change (e.g., Nielsen et al., 1992; Turner, Clancy, McQuade, & Cardenas, 1990).

Hamilton Rating Scale for Depression (HAM-D)

The Hamilton Rating Scale for Depression (Hamilton, 1960) represents a potentially valuable alternative to self-report questionnaires for the assessment of depressive symptom severity. Using this method, trained interviewers or clinicians make ratings of the presence and severity of specific symptoms of depression to derive a total score reflecting symptom severity. The HAM-D is almost certainly the most frequently employed observer-rated measure of depressive symptom severity.

Although the original version of the measure had 21 items, a 17-item version is the most commonly employed measure at the present time. Items are formatted as a checklist of symptoms with ratings of severity for each item ranging from either 0 to 4 or 0 to 2. Presumably, the decision to use only a 0-2 range was based on an assumption of the difficulty of making further discriminations in terms of severity for some symptoms. Range for each interviewer for the 17 items ranges from 0 to 50. A version of the scale with a manual for training interviewers was developed as part of the Early Clinical Drug Evaluation Program, and it is this version that is most commonly used (Zitman, Mennen, Griez, et al., 1990), although several others have published guidelines designed to enhance its reliability (Carroll, Feinburg, Smouse, et al., 1981; Potts, Daniels, Burnam et al., 1990; Reynolds & Kobak, 1990). Computerized versions have also been published (e.g., Kobak, Reynolds, & Rosenfeld, 1990), and the measure has been translated into

numerous languages. Relatively recently, a self-report measure based on the HAM-D has been developed, termed the Hamilton Depression Inventory (HDI; Reynolds & Kobak, 1995). The HAM-D usually takes between 15 and 20 minutes to administer.

Indices of internal consistency appear to vary considerably depending on the population and context. An international study yielded indices of only 0.48 before treatment, but 0.85 after treatment (Gaspar & Gilsdorf, 1990). Other published reports generally have yielded indices of greater than 0.80 when structured interview methods are employed (Potts et al., 1990). Indices of interrater agreement have also tended to be adequate, ranging from 0.65 (Maier, Phillip, & Heuser, 1988) to 0.90 (Hamilton, 1960; Rehm & O'Hara, 1985).

Indices of criterion-related validity have also tended to be very good. Correlations with global measures of depressive symptom severity have been reported to be in the 0.65 to 0.90 range, and correlations with other clinician ratings have typically been in the 0.80 to 0.90 range (Yonkers & Samson, 2000).

Like other measures of depressive symptom severity that include a high number of somatic items, concerns have been raised that the HAM-D may yield inflated rates of depressive disorder when employed in medical populations or the elderly in which a high prevalence of medical conditions is known to be present. Another concern that has been raised is the ability of the HAM-D to reliably discriminate depression from anxiety symptoms (Maier et al., 1988). The HAM-D has not been updated since prior to the publication of the DSM-III and this fact may limit its current sensitivity and specificity as a method for screening for the presence of depressive disorder or for reliable monitoring

of symptoms included in the current psychiatric nomenclature. Although the HAM-D has been encouraged because of its reliance on clinician interview and ratings rather than solely relying on respondent self-reports, data strongly suggest that its reliability can be improved with the use of a manual and adequate training of interviewers. Finally, the HAM-D has not been used extensively in the pain and pain management literatures and its sensitivity to change as a function of pain treatment has not been established.

Zung Self- Rating Depression Scale (Zung SDS)

The Zung Self-Rating Depression Scale (Zung, 1965) is another self-report measure of depressive symptom severity. The scale was developed to be short and simple to administer while at the same time including items reflecting the affective, cognitive, behavioral, and physiological components of depression. Items were specifically selected on the basis of diagnostic criteria at the time of its development and available factor analytic studies. The scale does not include several somatic symptoms acknowledged to be present in atypical depression including appetite and weight gain and hypersomnia. The original version of the Zung SDS included 20 items, with 10 items keyed in the positive direction and 10 in the negative direction to control for response bias. Respondents report frequency of occurrence of each symptom on four point scales. A revised version altered the original wording of two items, but the measure has largely remained the same since its original development and publication. The scale takes between 5 and 30 minutes to complete depending on the level of functioning of the respondent. The measure has been extensively translated and data supporting its strong

psychometric properties are available for many of these versions (Yonkers & Samson, 2000).

In one study of healthy volunteers, the index of internal consistency (Cronbach's alpha) was reported to be 0.79 and the split-half reliability coefficient was found to be 0.73. The criterion-related validity of the Zung SDS has been reported in several published studies, including correlations with other self-report (MMPI-Depression scale) and clinical interview measures (e.g., HAM-D) of depression ranging from 0.45 to 0.76 (e.g., Biggs et al., 1978). Although there have been several reports of the measure's sensitivity to change as a function of treatment for depression, a review of drug treatment studies found that the Zung SDS was specifically not sensitive to change relative to other depressive symptom severity measures (Moran & Lambert, 1983). There have also been significant challenges to its ability to yield reliable diagnoses of depression relative to other diagnostic categories (Guy, 1976). There are few reports of its use with chronic pain patients, and psychometric data to support its reliability and validity in this population are lacking.

Geriatric Depression Scale (GDS)

The Geriatric Depression Scale (GDS; Yesavage & Brink, 1983) was specifically developed to assess depressive symptom severity among elderly persons. The development of the instrument was encouraged by observations of the fact that all of the other self-report measures of the construct were developed and validated with medically healthy younger adults. These measures suffer from the criticism that they include

numerous somatic symptoms that are common among non-depressed elderly persons and that their format for responding may be difficult for some elderly persons.

The GDS consists of 30 “yes” versus “no” questions; 10 are negatively keyed and 20 are positively keyed. Questions are ordered with more “acceptable” items presented first. A shorter version of the measure has also been published that consists of 15 items (Sheikh & Yesavage, 1986). The total score for this version has been found to be highly correlated with the original version. An interview-based version has also been published that also is highly correlated with the original version and with the HAM-D (Jamison and Scogin, 1992). Finally, a telephone version has been demonstrated to have good agreement with the original version (Burke et al., 1995).

In the original publication, indices of internal consistency (0.94) and split-half reliability (0.94) were extremely high. These indices were significantly higher than those for the Zung SDS in the same sample. Correlations with the Zung SDS (0.84) and HAM-D (0.83) were also reported to be high, and the GDS was successful in discriminating mild from severe depressed groups in this same study. In this study, depressed elderly persons with arthritis were discriminated from non-depressed persons with arthritis. Brink et al. (1982) reported a high degree of sensitivity and specificity in discriminating depressed from non-depressed persons in a separate sample.

This measure seems to have substantial advantages for the assessment of depressive symptom severity among elderly persons. However, additional research with chronic pain samples will be necessary before its use in pain treatment outcome research can be supported.

Hospital Anxiety and Depression Scale (HADS)

The Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983) was specifically designed to screen for the presence of emotional distress among medically ill patients. In partial response to concerns raised about other measures of depression, in particular, the authors of this measure included only items that focus on the subjective experience of emotional distress, rather than physical signs. In addition, to further discriminate the experience of mood disturbance among medically, as opposed to psychiatrically, ill individuals, the depression subscale focuses on the experience of anhedonia, rather than on sadness.

The HADS is a self-report measure that includes only 14 items rated on 4-point Likert type scales. There are two subscales: depression and anxiety. Each subscale is comprised of 7 items. A test manual has been published (Snaith & Zigmond, 1994). The HADS has been translated into numerous languages.

Indices of internal consistency of the depression subscale have been reported to generally be above 0.90 (Moorey, Greer, Watson, et al., 1991; Zigmond & Snaith, 1983). Evidence of the criterion-related and discriminate validity of the HADS depression scale has also been reported. Advantages of the measure for the assessment of depressive symptom severity are its brevity and its development and standardization for medically ill, as opposed to psychiatric, populations.

West Haven-Yale Multidimensional Pain Inventory – Affective Distress scale (MPI-AD)

The West Haven- Yale Multidimensional Pain Inventory (MPI; Kerns, Turk & Rudy, 1985) has become one of the most commonly employed measures of psychosocial functioning in the field of pain. Among its several subscales is a three-item Affective Distress scale. Indices of internal consistency and factorial validity have been repeatedly found to be adequate. Although data on the utility of the Affective Distress subscale as a valid measure of depressive symptom severity is limited, the fact that the measure was developed specifically for the assessment of distress among persons with pain, the extensive experience with the measure in the pain field, and its brevity encouraged attention as a potentially useful measure of emotional distress among persons with pain. The fact that the MPI has been used extensively to evaluate outcome from pain interventions also encourages its use in pain intervention research. Several studies have reported on the sensitivity of the Affective Distress scale to change as a function of treatment, in particular (e.g., Marhold et al. 2001; McCarberg & Wolf, 1999).

Measures of anger

This section will provide only a cursory review of the available measures for the assessment of anger and the related construct of hostility, largely because of the absence of data supporting the relevance and utility of the measurement of anger in the context of pain treatment studies. As implied earlier, the failure to consider anger in the context of pain treatment is clearly not because of the rarity of anger or anger control problems among persons with chronic pain. On the contrary, as already noted, attention to the prevalence of anger among persons with chronic pain, and its potential role in the perpetuation, if not the development, of chronic pain and disability, is rapidly increasing

in the literature on the psychosocial aspects of pain. More likely, the failure to include anger as a target of pain treatment or to include measurement of anger as an important outcome of treatment rests with the dominance of historical attention on anxiety and depression, the absence of specific “anger” disorders in the psychiatric nomenclature, and the relative absence of efficacious treatments for excessive anger and anger control problems.

Fernandez (2002) provides a brief review of several different measures of anger, hostility, and anger expression that might have potential utility in the assessment of these variables among persons with clinical pain disorders. These include the following:

Buss Durkee Hostility Inventory (Buss & Durkey, 1957)

Overcontrolled Hostility Scale (Megargee, Cook, & Mendelsohn, 1967)

Hostility and Direction of Hostility Questionnaire (Caine, Foulds, & Hope, 1967)

Cook-Medley Hostility Scale (Cook & Medley, 1954)

Anger Self-Report (Zelin, Adler, & Myerson, 1972)

Reaction Inventory (Evans & Strangeland, 1971)

Anger Inventory (Novaco, 1974)

Multidimensional Anger Inventory (Siegal, 1986)

Targets and Reasons for Anger in Pain Sufferers (Fernandez, 1996)

State-Trait Anger Expression Inventory (Spielberger, 1988; 1999)

Among these measures, the later two on the list have found some utility in the assessment of anger, anger expression, and targets of anger among persons with chronic pain (Fernandez & Turk, 1995; Kerns et al., 1994; Okifuji, Turk, & Curran, 1999). As psychological and multidimensional interventions that emphasize reduction in anger

through assertiveness training, for example, or other similar intervention, incorporation of one or more of these measures may prove to be valuable in the assessment of outcomes (e.g., Keefe, Beupre, & Gil, 1996).

Multidimensional measures of emotional (psychological) functioning

Minnesota Multiphasic Personality Inventory (MMPI)

The Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1943) is by far the most commonly used objective measure of personality, and it is similarly the most commonly employed measure for the evaluation of psychological functioning of persons with pain. A recently revised version, known as the MMPI-2, is comprised of 567 true-false items that are used to derive scores on ten clinical scales, three validity scales, and fifteen new content scales (Hathaway, McKinley, Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989). The ten clinical scales are the most commonly examined scales in clinical settings. These scales are named:

Hypochondriasis, Depression, Hysteria, Psychopathic Deviate, Masculinity-Femininity, Paranoid, Psychasthenia, Schizophrenia, Mania, and Social Isolation. Respondent scores on these scales are converted into standard T-scores so that they may be compared to normative data. The revised version of the measure is thought to be more culturally sensitive and advantageous relative to the original version because the validation samples were more representative of the population of the US.

Nevertheless, significant concerns have been raised about the appropriateness of the either the MMPI or the MMPI-2 for use in the assessment of persons with chronic pain (Bradley & McKendree-Smith, 2001). Observed differences on the clinical scales

between pain and non-pain samples have been demonstrated to more likely reflect disease status rather than psychological functioning (e.g., Pincus, Callahan, Bradley, Vaughn, & Wolfe, 1986). An extensive research effort has focused on the identification of reliable subgroups of patients with chronic pain based on their MMPI profiles. The sum of this literature suggests that, although reliable subgroups can be identified, and despite evidence that the subgroups differ in terms of behavioral correlates of the experience of pain, it has yet to be demonstrated in a compelling fashion that the MMPI has value in characterizing patterns of coping with chronic pain over and above data derived from pain-specific measures (Bradley & Mckendree-Smith, 2001). This is particularly notable given the high response burden required for completion of the MMPI. In addition, inconsistent results from several studies challenge support for the value of the MMPI profiles as reliable predictors of pain treatment responsiveness (Guck, Meilman, Skultety, & Poloni, 1988; McCreary, 1985; Moore, Armentrout, Parker, & Kivlahan, 1986). Two recent studies stand in contrast to these relatively disappointing findings. In one study, Clark (1996) reported that the Negative Treatment Indicators (NTI) content scale from the MMPI-2 reliably predicted male patients' improvements in depressive symptom severity and physical capacity evaluations after multidisciplinary treatment. A study by Vendrig and colleagues demonstrated that scores on several MMPI-2 scales reliably predicted post-treatment changes on measures of pain intensity and disability (Vendrig, Derksen & deMey, 1999). Interestingly, in contrast to the Clark study findings, MMPI-2 scores did not predict post-treatment change on a similar measure of physical capacity. Similarly, results of studies designed to examine the sensitivity of the measure to change as a function of pain treatment have not been consistent or compelling.

Symptom Checklist – 90 Revised (SCL-90R)

The Symptom Checklist – 90 Revised (SCL-90R; Derogatis, 1983) requires respondents to rate the extent to which they have been bothered by each of 90 physical or mental health symptoms in the past week. Responses are used to derive nine specific standardized indices of psychological disturbance labeled as Somatization, Obsessive-Compulsive, Interpersonal Sensitivity, Depression, Anxiety, Hostility, Phobic Anxiety, Paranoid Ideation, and Psychoticism. A Global Severity Index may also be derived. The reliability and validity of the SCL-90R for the evaluation of psychiatric patients have been extensively reported in a manual for the instrument (Derogatis, 1983) and by others (Peveler & Fairburn, 1990).

Like the MMPI-2, the appropriateness of this measure for use in the assessment of persons with chronic pain has numerous critics. For example, Shutty, DeGood, and Schwartz (1986) were able to identify only five, rather than nine, reliable factors from the SCL-90R in a sample of chronic pain patients. These investigators also challenged the validity of evaluating patients with chronic pain by using norms developed from samples of psychiatric patients (Buckelew, DeGood, Schwartz, & Kerler, 1986). On the other hand, Jamison, Rock, and Parris (1988) identified three reliable subgroups of patients with chronic pain using the SCL-90R. These investigators demonstrated that patients with elevations on the subscales of the measure, relative to those with a profile consistent with normative data, reported significantly higher levels of disability, sleep disturbance, and emotional distress. Several other investigators have largely replicated these findings. Unfortunately, no data have been published in support of the ability of these subgroups or

the individual scales to predict pain treatment response (Bradley & McKendree-Smith, 2001). A 53-item version of the measure, the Brief Symptom Inventory (BSI; Derogatis & Spencer, 1983) has also been published, but its psychometric strengths have been the focus of little research in the chronic pain field. Finally, the sensitivity of the SCL-90R or the BSI to change as a function of treatment has not been adequately demonstrated.

Millon Behavioral Health Inventory (MBHI)

The Millon Behavioral Health Inventory (MBHI; Millon, Green, & Meagher, 1983) is a 150-item true-false measure specifically developed to evaluate the psychological functioning of persons with physical health problems. Numerous scales measure styles of patients' interactions with healthcare providers (e.g., Cooperative), major psychological stressors (Social Alienation), and response to treatments (e.g., Pain Treatment Responsivity) and illness (e.g., GI susceptibility). Adequate indices of reliability and validity have been reported, and the measure has clear advantages over either the MMPI or the SCL-90R since it was specifically designed and evaluated for use with physical health and illness populations. Despite these apparent advantages, results of studies designed to evaluate its predictive validity relative to treatment outcome evaluations have been discouraging (Gatchel, Deckel, Weinberg, & Smith, 1985; Gatchel, Mayer, Capra, Barnett & Diamond, 1986; Herron, Turner, Ersek, & Weiner, 1992; Sweet, Breuer, Hazlewood, Toye, & Pawl, 1985). Again, although the MBHI may have some utility in the characterization of persons with chronic pain, its utility as an outcome measure in pain treatment studies is not clear and demonstrations of its sensitivity to change as a function of treatment has not been forthcoming.

Illness Behavior Questionnaire (IBQ)

The Illness Behavior Questionnaire (IBQ; Pilowsky & Spence, 1975; 1994) is a 62-item true-false questionnaire designed to identify patterns of abnormal illness behavior. Seven scales are labeled: General Hypochondriasis, Disease Conviction, Psychological versus Somatic Focus of Disease, Affective Inhibition, Affective Disturbance, Denial of Life Problems Unrelated to Pain, and Irritability. Serious challenges to the reliability, factor structure, and validity of the IBQ have been raised, including concerns that it may be largely confounded by the respondent's level of anxiety or neurotic features (Bradley & McKendree-Smith, 2001). Dworkin and his colleagues, in a recent published report, suggested that elevated scores on the IBQ may reflect an appropriate, rather than an abnormal, response to chronic pain (Dworkin, Cooper, & Siegried, 1996). On the other hand, several investigators have reported data that appear to support the validity of at least some aspects of the IBQ as a measure of chronic illness (pain) behavior among chronic pain patients (e.g., Keefe, Crison, Maltbie, Bradley, & Gil, 1986; Waddell, Pilowsky, & Bond, 1989). Ultimately, concerns about the validity of this measure seem to outweigh its apparent strengths.

The Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36)

The SF-36 was developed as a general measure of perceived health status (Ware & Sherbourne, 1992). The measure is generally self-administered, although it has been used extensively in telephone administrations or in other interview settings. The measure contains 36 items that are combined to form eight scales: Physical Functioning, Physical

Role Functioning, Bodily Pain, General Health, Vitality, Social Functioning, Emotional Role Functioning, and Mental Health. Respondents use yes-no or 5 or 6 point scales to endorse the presence of degree of specific symptoms, problems, and concerns. The standard version of the measure employs a four-week recall period, but a more recent version uses a one-week timeframe. Scores on the scales range from 0-100 with higher scores indicating better health status and functioning. The measure takes about 10 to 15 minutes to complete.

The SF-36 has been extensively validated with large samples from the general population and across several demographic subgroups, including samples of healthy persons over 65 (Kazis, Miller, Clark, et al., 1998; Ware, Bayliss, Rogers, Kosinski, & Tarlov, 1996). A manual provides normative data for several medically ill groups (Ware, Snow, Kosinski, et al., 1993). Estimates of internal consistency (alphas) for most samples range from 0.62 to 0.94 for the subscales, with most estimates ranging over 0.80. Test-retest coefficients ranged from 0.43 to 0.81 for a six-month period, and from 0.60 to 0.81 for a two-week period (McHorney, Ware, Lu, et al., 1994). The SF-36 has been shown to correlate reasonably well with other criterion measures including the Sickness Impact Profile and the Duke Health Profile, measures of ability to work and utilization of healthcare resources, and other such clinically meaningful criteria such as “burden of care” (McHorney, Ware, Rczek, 1993; Ware, 2000). Factor analytic studies have supported the presence of two distinct factors labeled Physical Health and Mental Health Functioning that account for 82% of the measure’s variance (Kazis, Skinner, Rogers, Lee, Ren, & Miller, 1998; Ware, Kosinski, & Keller, 1994).

The SF-36 has only recently begun to be studied in chronic pain populations, including use as an outcome measure in pain intervention trials (Backonja, Beydoun, Edwards, et al., 1998; Katz, Harris, Larsen et al., 1992; Rowbotham, Harden, Stacey, et al., 1998). In one recent multi-site trial of cognitive-behavior therapy, exercise, and their combination, for persons with Gulf War Illness that included chronic, diffuse musculoskeletal pain as a primary feature, each of these treatments was found to be associated with improvements in the Mental Health Functioning component score (Donta et al., 2003). On a more negative note, Rogers and his colleagues reported that the SF-36 lacked reliability for the assessment of outcomes following multidisciplinary pain treatment and also questioned aspects of the measure's validity in discriminating dimensions of functional limitations (Rogers, Wittink, Wagner, Cynn, & Carr, 2000). Similar concerns about the sensitivity of the SF-36 to change have also been raised (McHorney & Tarlov, 1995). Continued examination of the sensitivity of the SF-36 Mental Health Functioning component to change as a function of pain interventions is indicated.

Profile of Mood States (POMS)

The POMS (McNair, Lorr & Droppleman, 1981; 1992) is a self-report instrument designed to assess six dimensions of mood: Tension-Anxiety (i.e., heightened musculoskeletal tension including reports of somatic tension and observable psychomotor manifestations of anxiety), Depression-Dejection (i.e., depression accompanied by a sense of personal inadequacy), Anger-Hostility (i.e., anger and antipathy toward others), Vigor-Activity (i.e., vigorousness, ebullience, and high energy), Fatigue-Inertia (i.e.,

weariness, inertia, and low energy level), and Confusion-Bewilderment (i.e., bewilderment, muddle-headedness appearing to be an organized-disorganized dimension of emotion). It is comprised of a list of 65 mood-related adjectives that requires respondents to report the degree to which each feeling or mood state has applied to them “for the past week, including today” using 0 (not at all) to 4 (extremely) Likert-type scales.

Reliability and validity of the measure were initially derived from a sample of persons presenting for healthcare at an academic medical center (n=1000). Persons who were illiterate, alcoholic, actively psychotic, and non-English speaking were excluded from the sample, and the age range was limited to those 60 years of age and under. Indices of internal consistency (alphas) for the six mood scales ranged from 0.84 for Confusion-Bewilderment to 0.95 for Depression-Dejection. Stability estimates (test-retest reliability correlations) ranged from 0.65 for Vigor-Activity to 0.74 for Depression-Dejection. Concurrent validity was examined via correlations with MMPI-2 scales. Correlations between scales of the POMS and analogous scales from the MMPI-2 were largely in the expected direction and significant, with coefficients ranging from -0.58 to 0.69. The POMS requires only about 3 to 5 minutes to administer.

The POMS has been used extensively in the pain treatment literature, and has been shown to be sensitive to change as a function of pain treatment (e.g., Backonja, et al., 1998; Rowbotham et al., 1998). Interestingly, however, its use has been largely limited to pharmaceutical trials, and it has yet to be employed in a large, randomized controlled trial of any psychological intervention.

Advantages of the POMS include its ease of administration, its brevity, its development on non-psychiatric populations, and its design to capture both negative and positive dimensions of emotional functioning. In particular, since the POMS has scales for anxiety, depression, and anger, three of the most important dimensions of emotional distress among persons with pain, the scale has an explicit advantage over any alternative scale. Within each of these three negative emotions, there is also an intuitive appeal to some of the items that are used to capture the construct for use with persons with chronic pain. For example, the Tension-Anxiety scale incorporates items that reflect both somatic and cognitive distress. The Depression-Dejection scale includes mood descriptors other than sadness that have been observed to be present in a large proportion of persons with chronic pain who meet criteria for current major depressive disorder but who otherwise deny feelings of depression. The inclusion of an Anger-Hostility scale is particularly novel and potentially an advantage of the POMS relative to any other comparable instrument. The Vigor-Activity scale represents a relatively unique opportunity to assess improvements in this key dimension of emotional functioning rather than relying on a reduction in negative mood and symptoms of emotional distress. The Fatigue-Inertia scale provides an opportunity to measure this common concomitant of the experience of chronic pain, especially when assessing pain treatment among persons with clinical pain conditions in which fatigue is particularly prevalent (e.g., pain in Multiple Sclerosis). The opportunity to attempt to discriminate effects of a pain intervention on fatigue and anergia, on the one hand, and other symptoms of emotional distress, on the other, may have particular utility in certain cases. Finally, given concerns about the

effects of certain pain medications on cognitive functioning, the Bewilderment-Confusion scale may also have some benefit.

Psychiatric Diagnostic Interviews

The use of structured psychiatric interviews is viewed as the state-of-the-art for reliable determination of the presence of psychiatric disorder. Having said this, unstructured clinical interviewing remains a more commonly used method for determining psychiatric diagnosis in the clinical setting (e.g., Sullivan, 2001), and even in most published clinical trials, the presence of a psychiatric diagnosis is generally not reported to be based on one of the more reliable methods for making this determination.

The two most commonly employed and widely researched psychiatric interviews are the Diagnostic Interview Schedule (DIS; Helzer & Robins, 1988; Robins, Helzer, Croughan, & Ratcliff, 1981) and the Structured Clinical Interview for DSM (SCID; Spitzer, Williams, Gibbon, & First, 1990). Neither the DIS nor the SCID has been used to examine psychiatric diagnosis among samples of persons with chronic pain. Even more important in the current context, neither measure has been used to examine effects of pain treatment on remission from psychiatric disorder, or even for examination of moderating effects of psychiatric disorder on pain treatment outcome. Nevertheless, these measures are briefly reviewed here because of their potential utility in characterizing pain treatment study samples, to control for psychiatric diagnosis in pain outcome studies, and for their potential utility, as yet unrealized, as reliable and valid measures of the efficacy of pain treatments as a contributor to remission from psychiatric disorder.

Diagnostic Interview Schedule (DIS)

The DIS, originally developed to provide reliable and valid diagnosis based on earlier versions of the Diagnostic and Statistical Manual of Mental Disorders (DSM), has been updated to correspond to the most recent edition (DSM-IV; American Psychiatric Association, 1994). Use of the DIS requires specialized training available at Washington University, home of its authors. Studies of the original version of the DIS revealed adequate test-retest stability estimates, with kappa coefficients ranging from .37 to .59 for lifetime psychiatric diagnosis over one-year periods (Helzer, Spitznagel & McEvoy, 1987; Vandiver & Sher, 1991). Kappa coefficients for diagnoses made by psychiatrists and lay interviewers ranged from .47 to 1.00 (Robins et al., 1981). Eaton and colleagues have provided evidence that the DIS may lead to underestimations of psychiatric diagnosis among the elderly, males, and those who have a relatively low level of impairment (Eaton, Neufeld, Chen, & Cai, 2000).

Structured Clinical Interview for DSM (SCID)

The SCID has also been recently updated to correspond to the DSM-IV (First, Spitzer, Williams, & Gibbon, 1997). Detailed materials are available to facilitate training in the use of this method. A computerized version of the measure has also been published (First, Gibbon, Williams, & Spitzer, 2000). Published rates of interrater agreement for primary psychiatric diagnosis hover around .70 (Williams, Gibbon, First, Spitzer, Davies, Borus, Howes, Kane, Pope, Rounsaville, & Wittchen, 1992).

Recommendations for the measurement of emotional distress in pain intervention research

The introductory sections of this monograph provided a brief review of a broad array of issues that should be taken into account in making decisions about the selection of measures for assessing the efficacy or effectiveness of pain interventions. To briefly summarize, the most salient issues and decisions include: how to reliably and validly assess emotional distress given its private, subjective, and complex nature; the importance of discriminating “levels” of analysis of emotional distress (i.e., mood, symptom clusters, disorders of emotional regulation) and dimensions of emotional distress (e.g., anxiety, depression, anger); and disagreements among clinicians and researchers alike about the nature of the relationship between pain and emotional distress. Perhaps as a function of the both the theoretical and empirical complexity of these issues and the lack of a consensus about any of them, it is not surprising that routine measurement of emotional distress has not been generally accepted as being of central importance in the assessment of pain interventions. On the other hand, the sheer prevalence of emotional distress among persons with clinical pain conditions, the pervasive negative effects of emotional distress on the experiences of pain and pain-related disability, the high costs associated with emotional distress among persons with pain, and the influence of emotional distress on pain treatment participation and outcomes all contribute to a view that emotional distress among persons with pain should be addressed in the context of any pain-relevant intervention, and should be assessed as an important, if not necessary, outcome of pain treatment.

Having reviewed these issues and agreed upon the importance of including reliable and valid measure(s) of emotional distress in pain treatment outcome research, it is equally clear that there is no current consensus on the appropriate targets for assessment of emotional distress, let alone their measurement. Particularly problematic is the fact that none of the most likely candidates for the assessment of emotional distress in pain intervention research were specifically developed for use in the assessment of persons with painful conditions. The few exceptions to this observation include measures that lack strong intuitive appeal as primary outcome measures of emotional distress due to their simplicity (e.g., MPI-Affective Distress scale), because they were not specifically designed to be used as outcome measures and likely lack sensitivity to change (Fear-Avoidance Beliefs Questionnaire), or have been only recently been proposed and lack sufficient reliability and validity data.

One final problem in the existing pain outcome literature that deserves serious attention is the apparent discrepancy between the pharmacological and non-pharmacological pain intervention literature in terms of the selection of measures of emotional distress. The POMS (McNair et al., 1971) is identified as the most commonly used measure of emotional distress in pharmacological treatment trials. This may be due to the simplicity of the measure and its low response burden, and perhaps most importantly, to the more general lack of specific attention to issues of comorbid psychiatric disorder, particularly depression. On the other hand, the inclusion of a measure of depressive symptom severity (e.g., the BDI or CES-D), and often a measure of level of anxiety (e.g., STAI), appears to represent the state-of-the-art for the assessment of outcomes from psychological pain treatments. Evaluations of the efficacy

or effectiveness of multidisciplinary pain treatment programs have more often included measures of depressive and/or anxiety symptom severity.

Ultimately, the inclusion of a multidimensional measure of emotional functioning, such as the POMS, as well as a measure of depressive symptom severity, such as the BDI or CES-D, is recommended for all pain intervention studies. The inclusion of a combination of these measures has distinct advantages over the selection of either one or the other of these measures alone. Support for this recommendation will highlight the known advantages of each of these measures, and will contrast their selection with alternative multidimensional and symptom specific measures. The use of a semistructured interview for psychiatric diagnosis, although it may be important in certain contexts, is probably unnecessary for most studies of pain outcome.

Among multidimensional measures of emotional functioning, the POMS appears to be the strongest candidate for inclusion in pain intervention research. The POMS has the distinct advantage of having been developed for the assessment of mood among non-psychiatric populations, and numerous psychometric studies have provided evidence to support its use in healthy and medically ill samples, including samples of persons with clinical pain conditions. It is simple to administer and has a particularly low response burden, requiring only 3 to 5 minutes for most persons to complete. As already emphasized, the POMS includes dimensions of emotional functioning that may provide the most comprehensive characterization of persons with chronic pain. The inclusion of positive, in addition to negative, dimensions of emotional functioning may prove to have advantages for pain treatment programs that are explicitly designed to promote wellness and adaptation, in addition to reduction in pain and disability, per se. Results of

numerous studies provide evidence of very good to excellent indices of reliability and stability of the measure. Published validity indices have largely been strong, and in particular, are based on the some of the best, and most contemporary, alternative measures of emotional functioning (e.g., the MMPI-2). The increasing use of the measure in pain outcome research, and demonstrations of sensitivity to change as a function of pain interventions, is a particularly compelling reason for this recommendation.

Limitations of the POMS are primarily two-fold. First, like all other multidimensional measures of emotional functioning are concerns about the intercorrelation of its subscales designed to measure discrete mood states. For most purposes, the use of a single composite score is indicated, and analyses focusing on change in the individual mood scales should be viewed with caution. Secondly, the measure is designed to serve as a measure of mood state, rather than as a more comprehensive measure of mood-related symptoms or disorder. For this reason, the importance of including a measure representative of a broader cluster of symptoms of emotional distress is strongly recommended.

As already described, the three most prevalent dimensions of emotional distress among persons with clinical pain disorders are anxiety, depression, and anger. Perhaps not surprisingly, assessment of anger and associated problems has not been a routine target in pain intervention research. Future research may help to identify the utility of targeting anger and its measurement in pain treatment research, but to date, there is not strong support for a recommendation to include its measurement as a routine component of pain intervention studies.

Anxiety, on the other hand, has an extensive history of dedicated attention in the pain field, including efforts to reduce this aspect of emotional distress in the context of pain treatment. For this reason, the pain research literature continues to focus on examination of the relationship between pain and anxiety, to refine pain treatments to further reduce pain-related anxiety, and to assess changes in anxiety as a function of treatment. Despite this relative emphasis, empirical research needs to continue to develop more specific measures of pain-related anxiety that may prove to be particularly beneficial in evaluating pain interventions. The most commonly used measure of anxiety in the pain intervention literature is by far the Spielberger State-Trait Anxiety Inventory (STAI). Unfortunately, this measure has few advantages over the POMS and likely lacks incremental utility in the assessment of emotional distress in pain intervention research. The STAI was initially developed for use with psychiatric populations, but fails to incorporate dimensions of anxiety disorders in its content. The correlation of its scores with other dimensions of emotional distress is known to be particularly high. The relative failure to employ the STAI in empirical trials of pain medications is also an important observation that raises concerns about its perceived relevance and its sensitivity as a measure of change. It is particularly noteworthy that the last decade has seen the emergence of a theory of pain-related fear and fear-avoidance and the proliferation of several alternative measures of pain-specific anxiety. Although none of these measures have been employed as indices of outcomes following pain treatment, it is likely that future research will incorporate such tests and may lead to recommendations for use of any one of these new measures as an important target for intervention and evaluation of its efficacy.

In contrast to these substantive concerns about inclusion of a measure of anxiety in pain intervention research, there is compelling evidence to support the selection of a measure of depressive symptom severity. The apparently high prevalence of discrete symptoms of depression (i.e., depressed or irritable mood, loss of interest in normally pleasurable activities, sleep dysfunction, anergia and fatigue, pervasive negative thinking, suicidal ideation) and of major depressive disorder and dysthymic disorder argue strongly for routine inclusion of a measure of depressive symptoms, as opposed to the sole reliance on a measure of depressed or dysphoric mood, in evaluation of pain treatment efficacy. Additional support comes from evidence of the analgesic potential of medications developed for the treatment of depressive disorder, observations that alleviation of depressed mood and other symptoms of depression may mediate the effectiveness of certain pain interventions, and evidence that depression may disrupt or interfere with successful pain treatment. Most importantly, if one agrees that the assessment of depression is particularly important in the assessment of pain intervention effectiveness, then the selection of a measure of the broader constellation of depressive symptoms is critical for a valid appraisal of the treatment's effects on this construct.

Among the several measures of depressive symptom severity, two have drawn the most attention from pain researchers and represent the state-of-the-art for assessment of this construct in the context of an evaluation of pain interventions. Based on a comprehensive review of the issues salient to making a recommendation for the adoption of a single measure of depressive symptom severity, if not an exhaustive review of the published literature, the BDI and CES-D clearly have the strongest support. Ultimately, the preponderance of the evidence led to the recommendation to employ the BDI.

A particularly important distinction between the BDI and the CES-D was the stated intent of the authors of these measures. Beck and colleagues specifically designed the BDI to be a reliable measure of depressive symptom severity **and** to assess change over time as a function of treatment. Indeed, the measure has a long and impressive history in this regard, and there is ample evidence to support its sensitivity to change as a function of both psychological and pharmacological treatments. As for the CES-D, it was designed to assess the level of depressive symptoms and to screen for the presence of depressive disorder in epidemiological studies of community, as opposed to clinical, samples. Although it has been employed extensively as a measure of change as a function of treatment, and it has ample evidence to support its sensitivity to change, this use has not been without its critics.

Both measures are simple to use and have a low response burden, although the BDI may take slightly longer to complete than the CES-D. Comparison of the evidence supporting the reliability and validity of each of these measures reveals few differences that can be upheld in support of one measure over the other. Both have evidence supporting their psychometric strengths across the broadest possible array of populations, including racial and ethnic minorities, women, and the elderly. Each measure has been challenged in terms of its inclusion of somatic symptoms that may inflate estimates of the prevalence of depressive disorder, although the preponderance of the evidence suggests that this is largely unfounded. Concerns have been raised about the sensitivity of the CES-D, in particular, in terms of its sensitivity and specificity as a tool for screening for depressive disorder, but even if this is the case, this concern should not be considered in an evaluation of its utility as a measure of outcome in pain intervention research.

There is over 25 years of research supporting the reliability and validity of the BDI and it is clearly the most extensively studied measure of emotional distress in the field of pain and pain management. The sheer volume of research on this instrument is the most compelling reason for its selection in this context.

Implications for future research

This review, although extensive, does not represent an exhaustive consideration of the available measures of pain-related emotional distress, let alone an empirically valid approach to the selection of measures of emotional distress for inclusion in pain intervention research. The conduct of a rigorous meta-analysis that could be employed to substantiate or dispute the recommendations of this monograph would be welcomed. Research designed to directly compare the reliability and sensitivity to change of the primary measures of emotional distress reviewed in this paper is also indicated. Similar studies have proved beneficial in evaluating the value of measures of pain-related disability and interference. Continued development and examination of measures of pain-related fear and fear-avoidance holds promise in advancing our understanding of the importance of these constructs, their potential for influencing refinements in pain interventions, and the measurement of this potentially important construct, particularly as a pain-specific alternative to more general measures of anxiety. Similarly, research on the construct of anger and its measurement is strongly encouraged. Finally, although it is likely that the routine inclusion of the POMS and BDI in pain intervention research will contribute substantially to our ability to aggregate findings across studies and sharpen our focus on the effects of pain interventions on emotional distress, continued development

of pain-specific measures of emotional distress, including depressive symptom severity, is encouraged.

References

- American Psychiatric Association (1994). Diagnostic and Statistical Manual of mental disorders (4th Edition). Washington, DC: American Psychiatric Association.
- Applebaum, K.A. et al. (1988). Cognitive-behavioral therapy of a veteran population with moderate to severe RA. Behavior Therapy, 19, 489-502.
- Asmundson, G.J., Jacobson, S.J., & Allardings, M.D. (1997). Fear and avoidance in dysfunctional chronic back pain patients. Pain, 69, 231-236.
- Asmundson, G.J., Jacobson, S.J., Allardings, M.D., & Norton, G.R. (1996). Social phobia in disabled workers with chronic musculoskeletal pain. Behaviour Research and Therapy, 34, 939-943.
- Atkinson, J.H., Slater, M.A., Patterson, T.L., Grant, I., & Garfin, S.R. (1991). Prevalence, onset, and risk of psychiatric disorders in chronic pain patients: A controlled study. Pain, 45, 111-121.
- Backonja, M., Beydoun, A., Edwards, K.R., Schwartz, S.L., Fonesca, V., Hes, M., LaMoreaux, L., Garofalo, E., for the Gabapentin Diabetic Neuropathy Study Group (1998). Gabapentin for the symptomatic treatment of painful neuropathy in patients with diabetes mellitus. Journal of the American Medical Association, 280, 1831-1836.
- Baker, D.G., Mendenhall, C.L., Simbartl, L.A., Magan, L.K., & Steinberg, J.L. (1997). Relationship between post-traumatic stress disorder and self-reported physical symptoms in Persian Gulf War Veterans. Archives of Internal Medicine, 157, 2076-2078.

- Banks, S.M. & Kerns, R.D. (1996). Explaining high rates of depression in chronic pain: A diathesis-stress framework. Psychological Bulletin, 119, 95-110.
- Beck, A.T. & Steer, R.A. (1993). Beck Depression Inventory. San Antonio, TX: Psychological Corporation.
- Beck, A.T., Steer, R.A., & Brown, G.K. (1986). Beck Depression Inventory-Second Edition Manual. San Antonio, TX: Psychological Corporation, Harcourt, Brace.
- Beck, A.T., Steer, R.A., Garbin, M.G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clinical Psychology Review, 8, 77-100.
- Beck, A.T., Ward, C.H., Mendelsohn, M., Mock, J. & Erbaugh, J. (1961). An inventory for measuring depression. Archives of General Psychiatry, 4, 561-571.
- Beckham, J.C., Crawford, A.L., Feldman, M.E., Kirby, A.C., Hertzberg, M.A., Davidson, J.R., & Moore, S.D. (1997). Chronic posttraumatic stress disorder and chronic pain in Vietnam combat veterans. Journal of Psychosomatic Research, 43, 379-389.
- Beitman, B.D., Mukerji, V., Kushner, M., Thomas, A.M., Russell, J.L., & Logue, M.B. (1991). Validating studies for panic disorder in patients with angiographically normal coronary arteries. Medical Clinics of North America, 75, 1143-1155.
- Beutler, L.E., Engle, D., Oro-Beutler, M.E. & Daldrup, R. (1986). Inability to express intense negative affect: A common link between depression and pain? Journal of Consulting and Clinical Psychology, 54, 752-759.
- Biggs, J.T., Wylie, L.T., Ziegler, V.E. (1978). Validity of the Zung Self-Rating Depression Scale. British Journal of Psychiatry, 132, 381-385.

- Boyd, J.H., Weissman, M.M., Thompson, W.D. et al. (1982). Screening for depression in a community sample: Understanding the discrepancies between depression symptom and diagnostic scales. Archives of General Psychiatry, *39*, 1195-1200.
- Bradley, L.A. & McKendree-Smith, N.L. (2001). Assessment of psychological status using interviews and self-report instruments. In D.C. Turk & R. Melzack (Eds.), Handbook of Pain Assessment, Second Edition. New York: Guilford Press.
- Bradley, L.A., Young, L.D., et al. (1987). Effects of psychological therapy on pain behavior of RA patients: Treatment outcome and six month follow-up. Arthritis and Rheumatism, *30*, 1105-1114.
- Breslau, N. (1985). Depressive symptoms, major depression, and generalized anxiety: A comparison of self-reports on CES-D and results from diagnostic interviews. Psychiatry Research, *15*, 219-229.
- Brink, T.A., Yesavage, J.A., Lum, O., et al. (1982). Screening tests for geriatric depression. Clinical Gerontologist, *1*, 37-43.
- Brown, F.F., Robinson, M.E., Riley, J.L., & Gremillion, H.A. (1996). Pain severity, negative affect, and microstressors as predictors of life interference in TMD patients. CRANIO, *14*, 63-70.
- Brown, G.K. (1990). A causal analysis of chronic pain and depression. Journal of Abnormal Psychology, *99*, 127-137.
- Buckalew, S.P., DeGood, D.E., Schwartz, D.P., & Kerler, R.M. (1986). Cognitive and somatic item response patterns of pain patients, psychiatric patients, and hospital employees. Journal of Clinical Psychology, *42*, 852-860.

- Buer, N. & Linton, S.J. (2002). Fear-avoidance beliefs and catastrophizing: Occurrence and risk factor in back pain and ADL in the general population. Pain, *99*, 485-491.
- Burke, W.J., Roccaforte, W.H., Wengel, S.P., et al. (1995). The reliability and validity of the Geriatric Depression Rating Scale administered by telephone. Journal of the American Geriatric Society, *43*, 674-679.
- Burns, J.W., Johnson, B.J., Mahoney, N., Devine, J., & Pawl, R. (1998). Cognitive and physical capacity process variables predict long-term outcome after treatment for chronic pain. Journal of Consulting and Clinical Psychology, *66*, 434-439.
- Buss, A.H. & Durkee, A. (1957). An inventory for assessing different kinds of hostility. Journal of Consulting Psychology, *21*, 343-349.
- Caine, T.M., Foulds, G.A., & Hope, K. (1967). Manual of the hostility and direction of hostility questionnaire. London: University of London.
- Carroll, B.J., Feinburg, M., Smouse, P.E., et al. (1981). The Carroll Rating Scale for Depression: Development, reliability, and validation. British Journal of Psychiatry, *138*, 194-200
- Catchlove, R.F.H. & Braha, R.E.D. (1985). A test to measure the awareness and expression of anger. Psychotherapy and Psychosomatics, *43*, 113-119.
- Chaves, J. & Barber, T.X. (1976). Hypnotism and surgical pain. In Mostovsky (Ed.), Behavior control and modification of psychological activity. Englewood Cliffs: Prentice Hall.

- Clark, M.E. (1996). MMPI-2 Negative Treatment Indicators Content and Content Component Scales: Clinical correlates and outcome prediction for men with chronic pain. Psychological Assessment, *8*, 32-38.
- Cook, W. & Medley, D. (1954). Proposed hostility and pharisaic-virtue scales for the MMPI. The Journal of Applied Psychology, *38*, 414-418.
- Crombez, G., Vlaeyen, J.W.S., Heuts, P.H.T.G., Lysens, R. (1999). Pain-related fear is more disabling than pain itself: Evidence on the role of pain-related fear in chronic back pain disability. Pain, *80*, 329-339.
- DeForge, B.R. & Sobal, J. (1988). Self-report depression scales in the elderly: The relationship between the CES-D and the Zung. International Journal of Psychiatry in Medicine, *18*, 325-328.
- DeGagne, T.A., Mikail, S.F., & D'Eon, J.L. (1995). Confirmatory factor analysis of a four-factor model of chronic pain evaluation. Pain, *60*, 195-202.
- Derogatis, L. (1983). The SCL-90R manual – II: Administration, scoring and procedures. Towson, MD: Clinical Psychometric Research.
- Derogatis, L. & Spencer, P. (1983). BSI Manual I: Administration and procedures. Baltimore, MD: Johns Hopkins University, Clinical Psychometric Unit.
- Donta, S.T., Clauw, D.J., Engel, C.C. et al. (2003). A randomized, multicenter trial of cognitive behavioral therapy and aerobic exercise for Gulf War Veterans' Illnesses: A Veterans Affairs Cooperative Study (CSP #470). Journal of the American Medical Association.
- Dworkin, R.H. & Caligor, E. (1988). Psychiatric diagnosis and chronic pain: DSM-III-R and beyond. Journal of Pain and Symptom Management, *3*, 87-98.

- Dworkin, R.H., Cooper, E.M., & Siegfried, R.N. (1996). Chronic pain and disease conviction. Clinical Journal of Pain, *12*, 111-117.
- Dworkin, R.H., Nagasako, E.M., Hetzel, R.D., & Farrar, J.T. (2001). Assessment of pain and pain-related quality of life in clinical trials. In D.C. Turk & R. Melzack (Eds.), Handbook of Pain Assessment, Second Edition. New York: Guilford Press.
- Eaton, W.W., Neufeld, K., Chen, L.S., & Cai, G. (2000). A comparison of self-report and clinical diagnostic interviews for depression: Diagnostic Interview Schedule and Schedules for Clinical Assessment in Neuropsychiatry in the Baltimore Epidemiologic Catchment Area follow-up. Archives of General Psychiatry, *57*, 217-222.
- Evans, D.R. & Strangeland, M. (1971). Development of the Reaction Inventory to measure anger. Psychological Reports, *29*, 412-414.
- Fechner-Bates, S., Coyne, J.C., & Schwenk, T.L. (1994). The relationship of self-reported distress to depressive disorders and other psychopathology. Journal of Consulting and Clinical Psychology, *62*, 550-559.
- Fernandez, E. (1996). The Targets and Reasons for Anger in Pain Sufferers: A structured interview. Unpublished manuscript, Southern Methodist University.
- Fernandez, E. (2002). Anxiety, depression, and anger in pain. Dallas, TX: Advanced Psychological Resources, Inc.
- Fernandez, E. & Turk, D.C. (1995). The scope and significance of anger in the experience of chronic pain. Pain, *61*, 165-175.

- Fishbain, D.A., Goldberg, M., Maegher, B.R., Steele, R., & Rosomoff, H. (1986). Male and female chronic pain patients categorized by DSM-III psychiatric diagnostic criteria. Pain, 26, 181-197.
- First, M.B., Gibbon, M., Williams, J.B.W., & Spitzer, R.L. (2000). SCID Screen Patient Questionnaire Computer Program. North Tonawanda, NY: Multi-Health Systems.
- First, M.B., Spitzer, R.L., Williams, J.B.W., & Gibbon, M. (1997). User's guide for the Structured Clinical Interview for the DSM-IV Axis I Disorders: SCID –I clinician version. Washington, DC: American Psychiatric Press.
- Flor, H., Birbaumer, N., & Turk, D.C. (1990). The psychobiology of chronic pain. Advances in Behavior Research and Therapy, 12, 47-84.
- Gaskin, M.E., Greene, A.F., Robinson, M.E., & Geisser, M.E. (1992). Negative affect and the experience of chronic pain. Journal of Psychosomatic Research, 36, 707-713.
- Gastpar, M & Gilsdorf, U. (1990). The Hamilton Depression Rating Scale in a WHO collaborative program. In P. Bech & A. Coppen (Eds.), The Hamilton Scales (Psychopharmacology Series 9). Berlin: Springer-Verlag.
- Gatchell R.J., Deckel A.W., Weinberg, N. & Smith, J.E. (1985). The utility of the Millon Behavioral Health Inventory in the study of chronic headaches. Headache, 25, 49-54.
- Gatchell, R.J., Mayer, T.G., Capra, P., Barnett, J., & Daimond, P. (1986). Millon Behavioral Health Inventory: Its utility in predicting physical function in patients

- with low back pain. Archives of Physical Medicine and Rehabilitation, 67, 878-882.
- Geisser, M.E., Roth, R.S., & Robinson, M.E. (1997). Assessing depression among persons with chronic pain using the Center for Epidemiological Studies – Depression Scale and the Beck Depression Inventory: A comparative analysis. Clinical Journal of Pain, 13, 163-170.
- Gerety, M.B., Williams, J.W., Mulrow, C.D. et al. (1994). Performance of case-finding tools for depression in the nursing home: Influence of clinical and functional characteristics and selection of optimal threshold scores. Journal of the American Geriatric Society, 42, 1103-1109.
- Gracely, R.H. (1992). Evaluation of multidimensional pain scales. Pain, 48, 297-300.
- Guck, T.P., Meilman, P.W., Skultety, M. & Poloni, I.D. (1988). Pain-patient Minnesota Multiphasic Personality Inventory subgroups: Evaluation of long-term treatment outcome. Journal of Behavioral Medicine, 11, 159-169.
- Guy, W. (1976). ECDEU Assessment Manual of Psychopharmacology-Revised (DHEW Publ No ADM 76-338). Rockville, MD: US Department of Health, Education, and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, NIMH Psychopharmacology Research Branch, Division of Extramural Research Programs.
- Hamilton, M. (1960). A rating scale for depression. Journal of Neurology, Neurosurgery, and Psychiatry, 23, 56-62.
- Hathaway, S.R. & McKinley, J. (1943). The Minnesota Multiphasic Personality Inventory. Minneapolis: University of Minnesota Press.

- Hathaway, S.R., McKinley, J.C., Butcher, J.N., Dahlstrom, W.G., Graham, J.R., Tellegen, A., & Kaemmer, B. (1989). Minnesota Multiphasic Personality Inventory-2: Manual for administration. Minneapolis: University of Minnesota Press.
- Haythornthwaite, J.A., Sieber, W.J. & Kerns, R.D. (1991). Depression and the chronic pain experience. Pain, 46, 177-184.
- Helzer, J.E. & Robins, L.N. (1988). The Diagnostic Interview Schedule: Its development, evaluation and use. Social Psychiatry and Psychiatric Epidemiology, 23, 6-16.
- Helzer, J.E., Spitznagel, E.L., & McEvoy, L. (1987). The predictive validity of lay Diagnostic Interview Schedule diagnoses in the general population: A comparison with physician examiners. Archives of General Psychiatry, 44, 1069-1077.
- Holroyd, K.A., Talbot, F., Holm, J.E., Pingel, J.D., Lake, A.E., & Saper, J.R. (1996). Assessing the dimensions of pain: A multitrait-multimethod evaluation of seven measures. Pain, 67, 259-265.
- Holroyd, K.A., Malinoski, P., Davis, M.K., & Lipchik, J.L. (1999). The three dimensions of headache impact: Pain, disability, and affective distress. Pain, 83, 571-578.
- Jamison, R.N., Rock, D.L., & Parris, W.C.V. (1988). Empirically derived Symptom Checklist 90 subgroups of chronic pain patients: A cluster analysis. Journal of Behavioral Medicine, 11, 147-158.

- Jamison, R.N., Rudy, T.E., Penzein, D.B., & Mosley, T.H. (1994). Cognitive-behavioral classifications of chronic pain: Replication and extension of empirically derived patient profiles. Pain, *57*, 277-292.
- Jamison, C. & Scogin, F. (1992). Development of an interview based geriatric depression rating scale. International Journal of Aging and Human Development, *35*, 193-204.
- Katon, W., Egan, K., & Miller, D. (1985). Chronic pain: Lifetime psychiatric diagnosis and family history. American Journal of Psychiatry, *142*, 1156-1160.
- Katz, J.N., Harris, T.M., Larson, M.G., et al. (1992). Predictors of functional outcomes after arthroscopic partial meniscectomy. Journal of Rheumatology, *19*, 1938-42.
- Kazis, L.E., Miller, D.R., Clark, J., et al. (1998). Health related quality of life in patients served by the Department of Veterans Affairs: results from the Veterans Health Study. Archives of Internal Medicine, *158*, 626-632.
- Kazis, L.E., Skinner, K., Rogers, W., Lee, A., Ren, X.S., & Miller, D. (1998). Health status and outcomes of veterans: physical and mental component summary scores (sf-36v). 1998 national survey of ambulatory care patients. Mid-year executive report. Office of Performance and Quality, Health Assessment Project, Center for Health Quality Outcomes and Economic Research, HSR&D Service, Veterans Administration, Washington, D.C. and Bedford (MA).
- Keefe, F.J., Beupre, P.M., & Gil, K.M. (1996). Group therapy for persons with chronic pain. In R.J. Gatchell & D.C. Turk (Eds.), Psychological approaches to pain management: A practitioner's handbook. New York: Guilford Press.

- Keefe, F.J., Crisson, J.E., Maltbie, A., Bradley, L.A., & Gil, K.M. (1986). Illness behavior as a predictor of pain and overt behavior patterns in chronic low back pain patients. Journal of Psychosomatic Research, *30*, 543-551.
- Kerns, R.D. (1996). Psychosocial factors: Primary or secondary outcomes? In J.N. Campbell & M.J. Mitchell (Eds.), Pain treatment centers at a crossroads: A practical conceptual reappraisal. Seattle: IASP Press.
- Kerns, R.D. & Haythornthwaite, J. (1988). Depression among chronic pain patients: Cognitive-behavioral analysis and rehabilitation outcome. Journal of Consulting and Clinical Psychology, *56*, 870-876.
- Kerns, R.D., Rosenberg, R. & Jacob, M.C. (1994). Anger expression and chronic pain. Journal of Behavioral Medicine, *17*, 57-67.
- Kerns, R.D., Turk, D.C., Holzman, A.D., & Rudy, T.E. (1986). Comparison of cognitive-behavioral and behavioral approaches to the outpatient treatment of chronic pain. The Clinical Journal of Pain, *1*, 195-203.
- Kerns, R.D., Turk, D.C., & Rudy, T.E. (1985). The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). Pain, *23*, 345-356.
- Khatami, M. & Rush, A.J. (1982). A one-year follow-up of the multimodal treatment for chronic pain. Pain, *14*, 45-52.
- Klapow, J.C., Slater, M.A., Patterson, T.L., Atkinson, J.H., Weickgenant, A.L., Grant, I., & Garfin, S.R. (1995). Psychosocial factors discriminate multidimensional clinical groups of chronic low back pain patients. Pain, *62*, 349-355.

Klapow, J.C., Slater, M.A., Patterson, T.L., Doctor, J.N., Atkinson, J.H., & Garfin, S.R.

(1993). An empirical evaluation of multidimensional clinical outcome in chronic low back pain patients. Pain, *55*, 107-118.

Kleinke, C.L. (1992). How chronic pain patients cope with pain: Relation to treatment outcome in a multidisciplinary pain clinic. Cognitive Therapy and Research, *16*, 669-685.

Klenerman, L., Slade, P.D., & Stanley, I.M. (1995). The prediction of chronicity in patients with an acute attack of low back pain in a general medical practice. Spine, *20*, 478-484.

Kobak, K.A., Reynolds, W.M., Rosenfeld, R., et al. (1990). Development and validation of a computer-administered version of the Hamilton Depression Rating Scale. Psychological Assessment, *2*, 56-63.

Koenig, H.G., George, L.K., Peterson, B.L., & Pieper, C.F. (1997). Depression in medically ill hospitalized older adults: Prevalence, characteristics, and course of symptoms according to six diagnostic schemes. American Journal of Psychiatry, *164*, 1376-1383.

Kori, S.H., Miller, R.P., & Todd, D.D. (1990). Kinesiophobia: A new view of chronic pain behavior. Pain Management, January/February, 35-43.

Kouyanou, K., Pither, C.E., Rabe-Hesketh, S., & Wessely, S. (1998). A comparative study of iatrogenesis, medication abuse, and psychiatric morbidity in chronic pain patients with and without medically explained symptoms. Pain, *76*, 417-426.

- Kulich, R.J., Mencher, P., Bertrand, C., & Maciewicz, R. (2000). Comorbidity of post-traumatic stress disorder and pain: Implications for clinical and forensic assessment. Current Review of Pain, *4*, 36-48.
- Larsen, D.K., Taylor, S., & Asmundson, G.J.G. (1997). Exploratory factor analysis of the Pain Anxiety Symptoms Scale in patients with chronic pain complaints. Pain, *69*, 27-34.
- Lethem, J., Slade, P.O., Troup, J.P.G., & Bentley, G. (1983). Outline of a fear-avoidance model of exaggerated pain perception. Behaviour Research and Therapy, *21*, 401-408.
- Linton, S.J. & Goettestam, K.G. (1985). Relations between pain, anxiety, mood, and muscle tension in chronic pain patients: A correlation study. Psychotherapy and Psychosomatics, *43*, 90-95.
- Magni, G., Caldieron, C., Rigatti-Luchini, S., & Merskey, H. (1990). Chronic musculoskeletal pain and depressive symptoms in the general population: An analysis of the 1st National Health and Nutrition Examination Survey data. Pain, *43*, 299-307.
- Maier, W., Phillip, M., Heuser, I. et al. (1988). Improving depression severity assessment, I: Content, concurrent, and external validity of three observer scales. Journal of Psychiatric Research, *22*, 3-12.
- Marhold, C., Linton, S.J., & Melin, L. (2001). A cognitive-behavioral return-to-work program: Effects on pain patients with a history of long-term versus short-term sick leave. Pain, *91*, 155-163.

- McCarberg, B. & Wolf, J. (1999). Chronic pain management in a health maintenance organization. Clinical Journal of Pain, *15*, 50-57.
- McCracken, L.M., Faber, S.D., & Janeck, A.S. (1998). Pain-related anxiety predicts nonspecific physical complaints in persons with chronic pain. Behaviour Research and Therapy, *34*, 927-933.
- McCracken, L.M., Gross, R.T., Aikens, J., & Carnrike, C.L.M. (1996). The assessment of anxiety and fear in persons with chronic pain: A comparison of instruments. Behaviour Research and Therapy, *34*, 927-933.
- McCracken, L.M., Gross, R.T., Sorg, P.J., & Edmands, T.A. (1993). Prediction of pain in patients with chronic low back pain: Effects of inaccurate prediction and pain-related anxiety. Behaviour Research and Therapy, *31*, 647-652.
- McCracken, L.M., Zayfert, C. & Gross, R.T. (1992). The Pain Anxiety Symptoms Scale: Development and validation of a scale to measure fear of pain. Pain, *50*, 67-73.
- McCreary, C. (1985). Empirically derived MMPI profile clusters and characteristics of low back pain patients. Journal of Consulting and Clinical Psychology, *53*, 558-560.
- McHorney, C. & Tarlov, A. (1995). Individual-patient monitoring in clinical practice: Are available health status surveys adequate. Quality of Life Research, *4*, 293-307.
- McHorney, C.A., Ware, J.E., Raczek, A.E. (1993). The MOS 36-item Short-Form Health Survey (SF-36), II: Psychometric and clinical tests of validity in measuring physical and mental health constructs. Medical Care, *31*, 247-263.

- McHorney, C.A., Ware, J.E., Lu, J.F.R. et al. (1994). The MOS 36-item Short-Form Health Survey (SF-36), III: Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Medical Care, *32*, 40-66.
- McNair, D.M., Lorr, M., & Doppleman, L.F. (1971). Profile of Mood States. San Diego, CA: Educational and Industrial Testing Service.
- McNeil, D.W. & Rainwater, III, A.J. (1998). Development of the Fear of Pain Questionnaire – III. Journal of Behavioral Medicine, *21*, 389-410.
- Megargee, E.I., Cook, P.E., & Mendelsohn, G.A. (1967). Development and validation of an MMPI scale of assaultiveness in overcontrolled individuals. Journal of Abnormal Psychology, *72*, 519-528.
- Mersky, H. (1978). Pain and personality. In R. Sternbach (Ed.), The psychology of pain. New York: Raven Press.
- Mikail, S.F., DuBreuil, S., & D'Eon, J.L. (1993). A comparative analysis of measures used in the assessment of chronic pain patients. Psychological Assessment, *5*, 117-120.
- Millon, T., Green, C. & Meagher, R. (1983). Millon Behavioral Health Inventory manual (3rd edition). Minneapolis, MN: National Computer Systems.
- Mongini, F., Defilippi, N., & Negro, C. (1997). Chronic daily headache. A clinical and psychological profile before and after treatment. Headache, *37*, 83-87.
- Moore, J.E., Armentrout, D.P., Parker, J.C., & Kivlahan, D.R. (1986). Empirically-derived pain-patient MMPI subgroups: Prediction of treatment outcome. Journal of Behavioral Medicine, *9*, 51-63.

- Moorey, S., Greer, S., Watson, M., et al. (1991). The factor structure and factor stability of the Hospital Anxiety and Depression Scale in patients with cancer. British Journal of Psychiatry, *158*, 255-259.
- Moran, P.W. & Lambert, M.J. (1983). A review of current assessment tools for monitoring changes in depression: In M.D. Lamber, E.R. Christensen and S.S. DeJolie (Eds.), The assessment of psychotherapy and outcomes. New York: Wiley.
- Myers, J.K. & Weissman, M.M. (1980). Use of a self-report symptom scale to detect depression in a community sample. American Journal of Psychiatry, *137*, 1081-1084.
- Naughton, M.J. & Wiklund, I. (1993). A critical review of dimension-specific measures of health-related quality of life in cross-cultural research. Quality of Life Research, *2*, 397-432.
- Nicholas, M.K., Wilson, P.H. & Goyen, J. (1991). Operant-behavioral and cognitive-behavioral treatment for chronic low back pain. Pain, *48*, 339-347.
- Nielsen, W.R., Walker, C. & McCain, G.A. (1992). Cognitive-behavioral treatment of fibromyalgia: Preliminary findings. The Journal of Rheumatology, *19*, 98-103.
- Novaco, R.W. (1974). The effect of disposition of anger and degree of provocation on self-report and physiological measures of anger in various modes of provocation. Unpublished manuscript, Indiana University, Bloomington, IN.
- Okifuji, A. & Turk, D.C. (2001). Assessment of treatment outcomes in clinical practice: A survival guide. In D.C. Turk & R. Melzack (Eds.), Handbook of Pain Assessment, Second Edition. New York: Guilford Press.

- Okifuji, A., Turk, D.C. & Curran, S.L. (1999). Anger in chronic pain: Investigations of anger targets and intensity. Journal of Psychosomatic Research, *47*, 1-12.
- Pfingsten, M., Kroner-Herwig, B., Leibing, E., Kronshage, U., & Hildebrandt, J. (2000). Validation of a German version of the Fear Avoidance Beliefs Questionnaire (FABQ). European Journal of Pain, *4*, 259-266.
- Pfingsten, M., Leibing, E., Harter, W., Kroner-Herwig, B., Hempel, D., Kronshage, U., & Hildebrandt, J. (2001). Fear-avoidance behavior and anticipation of pain in patients with chronic low back pain: A randomized controlled study. Pain Medicine, *2*, 259-266.
- Philips, H.C. (1987). The effects of behavioral treatment on chronic pain. Behaviour Research and Therapy, *25*, 365-377.
- Pilowsky, I. & Spence, N.D. (1975). Patterns of illness behavior in patients with intractable pain. Journal of Psychosomatic Research, *19*, 279-287.
- Pilowsky, I. & Spence, N.D. (1994). Manual for the Illness Behavior Questionnaire (IBQ). Unpublished manual (available from the authors).
- Pincus, T., Callahan, L.F., Bradley, L.A., Vaughn, W.K., & Wolfe, F. (1986). Elevated MMPI scores for hypochondriasis, depression, and hysteria in patients with rheumatoid arthritis reflect disease rather than psychological status. Arthritis and Rheumatism, *29*, 1456-1466.
- Polatin, P. B., Kinney, R. K., Gatchel, R. J., Lillo, E., & Mayer, T. G. (1993). Psychiatric illness and chronic low-back pain. The mind and the spine-- which goes first? Spine, *18*(1), 66-71.

- Potts, M.K., Daniels, M., Burnam, A. et al. (1990). A structured interview version of the Hamilton Depression Rating Scale: Evidence of reliability and versatility of administration. Journal of Psychiatric Research, 24, 335-350.
- Preveler, R.C. & Fairburn, C.G. (1990). Measurement of neurotic symptoms by self-report questionnaire: Validity of the SCL-90-R. Psychological Medicine, 20, 873-879.
- Radloff, L.S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. Applied Psychological Measurement, 1, 385-401.
- Radloff, L.S. & Locke, B.Z. (2000). Center for Epidemiologic Studies Depression Scale (CES-D). In Handbook of Psychiatric Measures. Washington, DC: American Psychiatric Association.
- Rehm, L. & O'Hara, M. (1985). Item characteristics of the Hamilton Rating Scale for Depression. Journal of Psychiatric Research, 19, 31-41.
- Reynolds, W.M. & Kobak, K.A. (1990). Reliability and validity of the Hamilton Depression Inventory: A paper and pencil version of the Hamilton Rating Scale Clinical Interview. Psychological Assessment, 7, 472-483.
- Richardson, I.H., Richardson, P.H., Williams, A.C. de C., Featherstone, J., & Harding, V.R. (1994). The effects of a cognitive-behavioural pain management programme on the quality of work and employment status of severely impaired chronic pain patients. Disability and Rehabilitation, 16, 26-34.
- Roberts, R.E. (1980). Reliability of the CES-D in different ethnic contexts. Psychiatry Research, 2, 125-134.

- Roberts, R.E. & Vernon, S.W. (1983). The Center for Epidemiological Studies-Depression Scale: Its use in a community sample. American Journal of Psychiatry, 140, 41-46.
- Robins, L.N., Helzer, J.E., Croughan, J., & Ratcliff, K.S. (1981). National Institute of Mental Health Diagnostic Interview Schedule. Archives of General Psychiatry, 38, 381-389.
- Robinson, M.E., Myers, C., Sadler, I.J., Riley, J.L., Kvaal, S.A., & Geisser, M.E. (1997). Bias effects in three common self-report assessment measures. Clinical Journal of Pain, 13, 74-81.
- Robinson, M.E. & Riley, III, J.L. (1999). The role of emotion in pain. In R.J. Gatchell & D.C. Turk (Eds.), Psychosocial factors in pain: Critical perspectives. New York: Guilford Press.
- Rodin, G. & Voshart, K. (1986). Depression in the medically ill: An overview. American Journal of Psychiatry, 143, 696-705.
- Rogers, W.H., Wittink, H., Wagner, A., Cynn, D. & Carr, D.B. (2000). Assessing individual outcomes during outpatient multidisciplinary chronic pain treatment by means of an augmented SF-36. Pain Medicine, 1, 44-54.
- Romano, J.M. & Turner, J.A. (1985). Chronic pain and depression: Does the evidence support a relationship? Psychological Bulletin, 97, 18-34.
- Rowbotham, M., Harden, N., Stacey, B., Bernstein, P., Magnus-Miller, L. for the Gabapentin Postherpetic Neuralgia Study Group (1998). Journal of the American Medical Association, 280, 1837-1842.

- Sheikh, J.I. & Yesavage, J.A. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. Clinical Gerontologist, 5, 165-173.
- Shutty, M.S., DeGood, D.E., & Schwartz, D.P. (1986). Psychological dimensions of distress in chronic pain patients: A factor analytic study of Symptom Checklist – 90 responses. Journal of Consulting and Clinical Psychology, 54, 836-842.
- Siegel, J.M. (1986). The Multidimensional Anger Inventory. Journal of Personality and Social Psychology, 51, 191-200.
- Snaith, R.P. & Zigmond, A.S. (1994). The Hospital Anxiety and Depression Scale Manual. Windsor, Berkshire, England: Nfer-Nelson.
- Spielberger, C.D. (1988). State-Trait Anger Expression Inventory Professional Manual. Odessa, FL: Psychological Assessment.
- Spielberger, C.D. (1999). The State-Trait Anger Expression Inventory-2 Professional Manual. Odessa, FL: Psychological Assessment.
- Spielberger, C.D., Gorsuch, R.L., & Lushene, R. (1970). Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press.
- Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). Manual for the State-Trait Anxiety Inventory (Form Y). Palo Alto: Consulting Psychologists Press.
- Spitzer, R.L., Williams, J.B.W., Gibbon, M., & First, M.B. (1990). Structured Clinical Interview for DSM-III-R. Washington, DC: American Psychiatric Press.

- Sullivan, M.D. (2001). Assessment of psychiatric disorders. In D.C. Turk & R. Melzack (Eds.), Handbook of Pain Assessment, Second Edition. New York: Guilford Press.
- Summers, J.D., Rapoff, M.A., Varghese, G., Porter, K., & Palmer, R.E. (1991). Psychosocial factors in chronic spinal cord injury pain. Pain, *47*, 183-189.
- Sweet, J.J., Breuer, S.R., Hazlewood, L.A., Toye, R., & Pawl, R.P. (1985). The Millon Behavioral Health Inventory: Concurrent and predictive validity in a pain treatment center. Journal of Behavioral Medicine, *8*, 215-226.
- Taenzer, P., Melzack, R., & Jeans, M.E. (1986). Influence of psychological factors on postoperative pain, mood, and analgesic requirements. Pain, *1986*, 331-342.
- Turk, D.C. (1996). Efficacy of multidisciplinary pain centers in the treatment of chronic pain. In M.J.M. Mitchell & J.N. Campbell (Eds.), Pain treatment centers at a crossroads: A practical and conceptual reappraisal. Seattle, WA: IASP Press.
- Turk, D.C. (2002). Clinical effectiveness and cost-effectiveness of treatments for patients with chronic pain. The Clinical Journal of Pain, *18*, 355-365.
- Turk, D.C., Meichenbaum, D., & Genest, M. (1983). Pain and behavioral medicine: A cognitive-behavioral perspective. New York: Guilford Press.
- Turk, D.C. & Melzack, R. (2001). Handbook of Pain Assessment, Second Edition. New York: Guilford Press.
- Turk, D.C. & Okifuji, A. (1994). Detecting depression in chronic pain patients: Adequacy of self-reports. Behaviour Research and Therapy, *32*, 9-16.

- Turk, D.C. & Rudy, T.E. (1988). Toward an empirically derived taxonomy of chronic pain patients: Integration of psychological assessment data. Journal of Consulting and Clinical Psychology, *56*, 233-238.
- Turk, D.C. & Rudy, T.E. (1990). The robustness of an empirically derived taxonomy of chronic pain patients. Pain, *43*, 27-35.
- Turner, J.A. (1982). Comparison of group progressive-relaxation and cognitive-behavioral group therapy for chronic low back pain. Journal of Consulting and Clinical Psychology, *50*, 757-765.
- Turner, J.A., Clancy, S., McQuade, K.J., & Cardenas, D.D. (1990). Effectiveness of behavioral therapy for chronic low back pain: A component analysis. Journal of Consulting and Clinical Psychology, *58*, 573-579.
- Vandiver, T. & Sher, K.J. (1991). Temporal stability of the Diagnostic Interview Schedule. Psychological Assessment, *3*, 277-281.
- Vendrig, A.A., Derksen, J.L., & deMey, H.R. (1999). Utility of selected MMPI-2 scales in the outcome prediction for patients with chronic back pain. Psychological Assessment, *11*, 381-385.
- Von Korff, M. & Simon, G. (1996). The relationship between pain and depression. British Journal of Psychiatry, *168* (Supplement 30), 101-108.
- Waddell, G., Newton, M., Henderson, I., Somerville, D., & Main, C. (1993). A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain, *52*, 157-168.
- Waddell, G. Pilowsky, I., & Bond, M.R. (1989). Clinical assessment and interpretation of abnormal illness behavior in low back pain. Pain, *39*, 41-53.

- Wade, J.B., Price, D.D., Hamer, R.M., & Schwartz, S.M. (1990). An emotional component analysis of chronic pain. *Pain*, *40*, 303-310.
- Ware, J.E., Bayliss, M.S., Rogers, W.H., Kosinski, M., & Tarlov, A.R. (1996). Differences in 4-year health outcomes for elderly and poor, chronically ill patients treated in HMO and fee-for-service systems. *Journal of the American Medical Association*, *276*, 1039-1047.
- Ware, J.E., Kosinski, M., & Keller, S.D. (1994). *SF-36 physical and mental health summary scales: a users manual*. Boston (MA): Health Assessment Lab, New England Medical Center.
- Ware, J.E. & Sherbourne, C.D. (1992). The MOS 36-item Short-Form Health Survey (SF-36), I: Conceptual framework and item selection. *Medical Care*, *30*, 473-483.
- Weissman, M.M., Sholomskas, D., Portenger, M., et al. (1977). Assessing depressive symptoms in five psychiatric populations: A validation study. *American Journal of Epidemiology*, *106*, 203-214.
- Williams, J.B., Gibbon, M., First, M.B., Spitzer, R.L., Davies, M.J., Borus, M., Howes, J., Kane, J., Pope, H.G., Rounsaville, B., & Wittchen, H-V. (1992). The Structured Clinical Interview for DSM-III-R (SCID) II: Multisite test-retest reliability. *Archives of General Psychiatry*, *49*, 630-636.
- Williams, A.C. deC., Richardson, P.H., Nicholas, M.K., Pither, C.E., Harding, V.R., Ridout, K.L., Ralphs, J.A., Richardson, I.H., Justins, D.M. & Chamberlain, J.H. (1996). Inpatient vs. outpatient pain management: Results of a randomized controlled trial. *Pain*, *66*, 13-22.

- Yesavage, J.A. & Brink, T.L. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. Journal of Psychiatric Research, *17*, 37-49.
- Yonkers, K.A. & Samson, J. (2000). Mood disorders measures. In American Psychiatric Association, Handbook of Psychiatric Measures. Washington, DC: American Psychiatric Association.
- Zelin, M.L., Adler, G., & Myerson, P.G. (1972). Anger self-report: An objective questionnaire for the measurement of aggression. Journal of Consulting and Clinical Psychology, *39*, 340.
- Zigmond, A.S. & Snaith, R.P. (1983). The Hospital Anxiety and Depression Scale. Acta Psychiatrica Scandinavia, *67*, 361-370.
- Zitman, F.G., Mennen, M.F.G., Griez, E. et al. (1990). The different versions of the Hamilton Depression Rating Scale. In P. Bech & A. Coppen (Eds.), The Hamilton Scales (Psychopharmacology Series 9). Berlin: Springer-Verlag
- Zung, W.W.K. (1965). A self-rating depression scale. Archives of General Psychiatry, *12*, 63-65.