



Prospective Comparison of the Minnesota Multiphasic Personality Inventory-2 (MMPI-2) and MMPI-2-Restructured Form (MMPI-2-RF) in Predicting Treatment Outcomes Among Patients with Chronic Low Back Pain

Anthony M. Tarescavage¹ · Judith Scheman² · Yossef S. Ben-Porath³

Published online: 15 February 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

The purpose of the current study was to examine the relative utility of the most updated MMPI adult instrument, the Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF), which was designed to address psychometric limitations of the MMPI-2. To this end, we compared mean scores and correlates of emotional distress treatment outcomes using the Depression Anxiety Stress Scales in a sample of 230 patients (73 males, 157 females) who had completed an interdisciplinary chronic pain rehabilitation program. Structural equation modeling analyses indicated that higher scale scores from all the MMPI-2-RF substantive domains were meaningfully associated with worse emotional distress outcomes, whereas the MMPI-2 Clinical Scales generally did not have any meaningful associations. Similar results were found in additional analyses using a clinically significant change framework with more direct clinical implications. The results of this study provide preliminary support for the use of the MMPI-2-RF among patients with chronic low back pain.

Keywords MMPI-2-RF · Chronic pain · Back pain · Treatment outcome · Applied assessment

Introduction

Chronic low back pain has a considerable impact on society. The prevalence of the disorder is rising (Freburger et al., 2009; Rubin, 2007), which is a prominent cause of disability (McNeil & Binette, 2001) and sick days (LaBar, 1992), and it has substantial economic influence (Guo, Tanaka, Halperin, & Cameron, 1999; Katz, 2006). Moreover, it is associated with psychological problems, as the 12-month prevalence rates of mood and anxiety disorders in this population are 17.5 and 26.5% (Von Korff et al., 2005), respectively,

which are nearly double the general population prevalence rates (Kessler et al., 2004). Chronic pain is also associated with the increased rates of illicit drug use, particularly opioid abuse (Manchikanti et al., 2006).

According to the biopsychosocial perspective of pain (Gatchel, McGeary, McGeary, & Lippe, 2014; Gatchel, Peng, Peters, Fuchs, & Turk, 2007), biological, psychological, and social factors interact to influence the experience of pain. Gatchel et al. (2007) provide an overview of how these factors affect the perception of illness, noting that pertinent psychological factors include mood problems, such as anxiety and depression, as well as cognitions that may lead to pain catastrophizing. The American College of Physicians and the American Pain Society recommend interdisciplinary treatment with an assessment of these and other psychosocial factors (Chou et al., 2007). They have been found to be stronger predictors of outcome than physical examinations, severity of pain, and duration of pain (Chou et al., 2007).

Psychological testing is one way to assess for these factors, with the Minnesota Multiphasic Personality Inventory (MMPI) (Hathaway & McKinley, 1943) and MMPI-2 (Butcher et al., 2001) historically having been the most

✉ Anthony M. Tarescavage
atarescavage@jcu.edu

¹ Department of Psychological Sciences, John Carroll University, 1 John Carroll Boulevard, University Heights, OH 44118, USA

² Digestive Disease and Surgical Institute, Cleveland Clinic Foundation, 9500 Euclid Avenue, Cleveland, OH 44195, USA

³ Department of Psychological Sciences, Kent State University, 144 Kent Hall, Kent, OH 44242, USA

frequently used psychological tests among chronic pain patients (Piotrowski, 1998; Piotrowski & Lubin, 1990). However, use of these instruments began to decline in chronic pain settings in the mid-to-late-1990s. During this time, a series of articles debating the utility of the instrument were published in *Pain Forum*. Main and Spanswick (1995) began the debate with an article entitled “Personality Assessment and the Minnesota Multiphasic Personality Inventory: 50 years on: Do we still need our security blanket?” The authors criticized the test for its psychometric shortcomings, writing, “Its inherent structural weaknesses undermine its clinical validity, even when it does provide additional clinical information” (p. 92). They called for prospective chronic pain outcome studies using advanced quantitative analyses such as structural equation modeling and measures “which reflect the world of pain rather than promulgate the sort of psychoarcheology represented by the MMPI and MMPI-2” (p. 95). Most of these concerns were echoed by other authors in the debate (Keefe, Lefebvre, & Beaupre, 1995; Turk & Fernandez, 1995). However, Bradley (1995) countered these claims by reviewing a series of research studies indicating that individuals can be reliably categorized into MMPI Scale score subgroups, which demonstrate concurrent associations with factors that may predict outcome (such as pain intensity, medication use, disability, and work status). Overall, most of the authors in the series agreed that significant problems with the test’s Clinical Scales (which were nearly identical to the MMPI’s Clinical Scales) limited the test’s utility in this setting.

Several years after the debate, the MMPI-2-Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008/2011), was released as an updated version of the MMPI-2. The MMPI-2-RF is a 338-item broadband measure of psychopathology with 51 scales. The nine Validity Scales of the test are designed to assess for problematic test-taking approaches, which include random and acquiescent responding, as well as over- and underreporting of psychological problems. The test’s substantive scales measure psychological constructs and are anchored by the nine Restructured Clinical (RC) Scales. The primary goal of the RC Scales project was to address the psychometric limitations of the Clinical Scales by substantially reducing the scale overlap and heterogeneity that complicated their interpretation and use in research, while still measuring the major distinctive core constructs assessed by each scale. The constructs measured by the scales were also more clearly tied to modern psychopathology models and constructs (Sellbom, Ben-Porath, & Bagby, 2008). These revisions address some of the primary concerns with the Clinical Scales advanced by authors in the debate.

The MMPI-2-RF test authors used similar modern scale development strategies for two substantive scale sets that complement the RC Scales: (1) the three Higher-Order

Scales that measure internalizing dysfunction, thought dysfunction, and externalizing dysfunction, broadly defined, and; (2) the 23 Specific Problems Scales that measure RC Scale subdomains or other, more narrowly focused constructs that are related to, but distinct from those measured by the RC Scales. Revised and improved versions of the MMPI-2 PSY-5 Scales, which measure broad domains of abnormal personality, are also included on the test. Overall, the MMPI-2-RF measures five substantive domains of personality and psychopathology: (1) Emotional Dysfunction; (2) Thought Dysfunction; (3) Behavioral/Externalizing Dysfunction; (4) Somatic/Cognitive Problems; and (5) Interpersonal Functioning (see Table 1 for scale descriptions).

McCord and Drerup (2011) demonstrated the improved interpretive utility of the RC Scales in comparison to the Clinical Scales in a chronic pain sample. These authors categorized 316 chronic pain patients into depressed and non-depressed diagnostic groups. The depression group included individuals diagnosed with major depression, dysthymia, and adjustment disorder, whereas the nondepressed group was not diagnosed with any form of mood disturbance. They compared mean scores on the Clinical and RC Scales across the two groups. In the nondepressed group, mean Clinical Scale elevations (i.e., scores $\geq 65T$) were found on scales 1 (Hypochondriasis), 2 (Depression), 3 (Hysteria), and 8 (Schizophrenia), whereas only RC1 (Somatic Complaints) produced a mean RC Scale elevation. In the depressed group, mean clinical elevations were observed for the following Clinical Scales: 1 (Hypochondriasis), 2 (Depression), 3 (Hysteria), 4 (Psychopathic Deviate), 6 (Paranoia), 7 (Psychastenia), and 8 (Schizophrenia). The pattern of elevations was consistent with the neurotic-triad cluster and code type typically found in Clinical Scale research in this setting, with prominent elevations on scales 1, 2, and 3. In stark contrast to the Clinical Scale findings, mean RC scale elevations were observed in the depressed group for only RCd (Demoralization), RC1 (Somatic Complaints), and RC2 (Low Positive Emotions), demonstrating substantially improved discriminant validity. McCord and Drerup (2011) summarize the implications of the findings from the depressed group:

“The clinician relying on the Clinical Scales would see clinical-range elevations on all scales except Scale 9, with extreme elevations on Scales 1, 2, and 3 and troubling elevations on 7 and 8 as well. In contrast, the RC Scales indicate three things: (a) a significant level of demoralization; (b) significant somatic complaints; and (c) depression. The latter set of data is far more consistent with the clinical diagnoses in the patient charts” (p. 145).

Current Study

Despite the substantial psychometric and interpretive improvements compared to the Clinical Scales, no study has

Table 1 Minnesota Multiphasic Personality Inventory-2-Restructured Form Scales

Validity Scales	
Inconsistent responding	
VRIN-r	Variable response inconsistency-random responding
TRIN-r	True response inconsistency-fixed responding
Overreporting	
F-r	Infrequent responses—responses infrequent in the general population
Fp-r	Infrequent psychopathology responses—responses infrequent in psychiatric populations
F _s	Infrequent somatic responses—somatic complaints infrequent in medical patient populations
FBS-r	Symptom validity—somatic and cognitive complaints associated at high levels with overreporting
RBS	Response bias scale—exaggerated memory complaints
Underreporting	
L-r	Uncommon virtues—rarely claimed moral attributes or activities
K-r	Adjustment validity—avowals of good psychological adjustment associated at high levels with underreporting
Higher-Order (H-O) Scales	
EID	Emotional/internalizing dysfunction—problems associated with mood and affect
THD	Thought dysfunction—problems associated with disordered thinking
BXD	Behavioral/externalizing dysfunction—problems associated with under-controlled behavior
Restructured Clinical (RC) Scales	
RCd	Demoralization—general unhappiness and dissatisfaction
RC1	Somatic complaints—diffuse physical health complaints
RC2	Low positive emotions—lack of positive emotional responsiveness
RC3	Cynicism—non-self-referential beliefs expressing distrust and a generally low opinion of others
RC4	Antisocial behavior—rule breaking and irresponsible behavior
RC6	Ideas of persecution—self-referential beliefs that others pose a threat
RC7	Dysfunctional negative emotions—maladaptive anxiety, anger, and irritability
RC8	Aberrant experiences—unusual perceptions or thoughts
RC9	Hypomanic activation—overactivation, aggression, impulsivity, and grandiosity
Specific Problem (SP) Scales	
Somatic/Cognitive Scales	
MLS	Malaise—overall sense of physical debilitation, poor health
GIC	Gastrointestinal complaints—nausea, recurring upset stomach, and poor appetite
HPC	Head pain complaints—head and neck pain
NUC	Neurological complaints—dizziness, weakness, paralysis, loss of balance, etc
COG	Cognitive complaints—memory problems, difficulties concentrating
Internalizing Scales	
SUI	Suicidal/death ideation—direct reports of suicidal ideation and recent suicide attempts
HLP	Helplessness/hopelessness—belief that goals cannot be reached or problems solved
SFD	Self-doubt—lack of confidence, feelings of uselessness
NFC	Inefficacy—belief that one is indecisive and ineffectual
STW	Stress/worry—preoccupation with disappointments, difficulty with time pressure
AXY	Anxiety—pervasive anxiety, frights, frequent nightmares
ANP	Anger proneness—becoming easily angered, impatient with others
BRF	Behavior-restricting fears—fears that significantly inhibit normal activities
MSF	Multiple specific fears—fears of blood, fire, thunder, etc
Externalizing Scales	
JCP	Juvenile conduct problems—difficulties at school and at home, stealing
SUB	Substance abuse—current and past misuse of alcohol and drugs
AGG	Aggression—physically aggressive, violent behavior
ACT	Activation—heightened excitation and energy level

Table 1 (continued)

Interpersonal Scales	
FML	Family problems—conflictual family relationships
IPP	Interpersonal passivity—being unassertive and submissive
SAV	Social avoidance—avoiding or not enjoying social events
SHY	Shyness—bashful, prone to feel inhibited and anxious around others
DSF	Disaffiliativeness—disliking people and being around them
Interest Scales	
AES	Aesthetic-literary interests—literature, music, the theater
MEC	Mechanical-physical interests—fixing and building things, the outdoors, sports
Personality Psychopathology Five (PSY-5) Scales	
AGGR-r	Aggressiveness-revised—instrumental, goal directed aggression
PSYC-r	Psychoticism-revised—disconnection from reality
DISC-r	Disconstraint-revised—under-controlled behavior
NEGE-r	Negative emotionality/neuroticism-revised—anxiety, insecurity, worry, and fear
INTR-r	Introversive/low positive emotionality-revised—social disengagement and anhedonia

From *MMPI-2-RF Manual for Administration, Scoring, and Interpretation* by Yossef S. BenPorath and Auke Tellegen. Copyright© 2008, 2011 by the Regents of the University of Minnesota. Reproduced by permission of the University of Minnesota Press. All rights reserved. “Minnesota Multiphasic Personality Inventory-2-RF®” and “MMPI-2-RF®” are trademarks owned by the Regents of the University of Minnesota

investigated use of the RC Scales to predict outcomes among chronic pain patients undergoing conventional conservative treatments. The purpose of the current study was to investigate the ability of the MMPI-2-RF to predict self-reported emotional distress outcomes among patients with chronic low back pain completing short-term interdisciplinary rehabilitation treatment. Because this is the first comprehensive investigation of the revised inventory, rather than test-specific hypotheses, in the context of discovery (Reichenbach, 1938), we investigated the association between all MMPI-2-RF scale scores with emotional distress outcomes (as measured by the Depression Anxiety Stress Scales) (Lovibond & Lovibond, 1995) after controlling for age and gender, as well as pain intensity and duration of pain. In line with suggestions by Main and Spanswick (1995), we examined these associations using structural equation modeling. Finally, we compared the scores of interpretive utility and the predictive capacity of the MMPI-2-RF scales and the MMPI-2 Clinical Scales, and we expected that the MMPI-2-RF scales would demonstrate substantially greater interpretive utility and larger effect sizes in predicting outcome given its structural and theoretical improvements.

Methods

Participants

Participants were drawn from an archival sample of 278 non-consecutive chronic pain patients (93 males, 185 females) who presented with lower back pain to a 3–4 week interdisciplinary pain treatment program in Northeast Ohio and were

administered the MMPI-2 as well as the Depression Anxiety Stress Scales (DASS) at intake. Overall, 249 (89.6%) of these individuals were eligible for inclusion because they completed the program and were administered the DASS at discharge. Completers participated in the program for an average of 20.6 days ($SD = 6.8$).

MMPI-2 items were used to calculate MMPI-2-RF scale scores, which is possible because all 338 MMPI-2-RF items are included in the MMPI-2 booklet. Past research has demonstrated the relative comparability of MMPI-2-RF scale scores generated from both booklets (Tellegen & Ben-Porath, 2008/2011; Van der Heijden, Egger, & Derksen, 2010). An additional 19 participants were excluded from the analyses because they produced invalid MMPI-2-RF profiles according to the test authors’ published guidelines, which included cannot say, $CNS \geq$ Raw score 18; variable response inconsistency, $VRIN-r \geq 80$; true response inconsistency, $TRIN-r \geq 80$; infrequent responding, $F-r = 120$; and infrequent psychopathology responses, $Fp-r \geq 100$ (Ben-Porath & Tellegen, 2008/2011).

The final sample included 230 patients (73 males, 157 females) after exclusions.¹ The majority of the sample was married (62.6%) and other marital statuses included never

¹ There were a total of 14 additional participants (6% of sample) who would have been excluded from the study if we removed individuals with elevated scores ($\geq 100T$) on Fs, FBS-r, or RBS. According to Ben-Porath and Tellegen (2008/2011), elevated scores on these three validity scales suggest scores on the Somatic/Cognitive scales may be invalid. We ran the study analyses after excluding these 14 individuals, and the interpretation of the findings did not meaningfully change.

married (19.6%), divorced (10.0%), widowed (3.0%), and separated (1.3%). The average age was 46.5 ($SD = 14.5$), and the average years of education was 14.3 ($SD = 3.1$). No significant differences were observed on most demographic variables between the excluded individuals and the final sample (p 's > .23). However, excluded individuals were more likely to have a marital status of separated, $\chi^2(5) = 11.455$, $p = .043$, Std. residual = 2.6. In the final sample, the most common DSM-IV-TR (American Psychiatric Association, 2000) diagnoses included major depressive disorder (48.8%), a substance use disorder (36.3%), an anxiety disorder (18.2%), a bipolar disorder (4.9%), post-traumatic stress disorder (2.6%), or a somatization disorder (0.9%)(categories are not mutually exclusive). To some extent, these rates may underestimate comorbid psychopathology because billing was not based the presence of a mental disorder. Common comorbid medical conditions included joint pain (44.8%), neck pain (38.7%), foot pain (29.1%), fibromyalgia (25.8%), migraine (19.7%), arthritis (18.3%), neuropathic pain (17.8%), abdominal pain (10.4%), myofascial pain (5.7%), chronic regional pain syndrome (5.2%), chronic fatigue syndrome (3.9%), tension headache (3.5%), dizziness (3.1%), and diabetes (3.1%). In terms of medications at intake, 67.6% were prescribed an antidepressant and 61.9% were prescribed pain medications.

Measures

MMPI-2-RF

The MMPI-2-RF is described in detail in the Introduction. Psychometric properties of scores from the instrument are reported by Tellegen and Ben-Porath (2008/2011). Descriptive statistics and reliability estimates are provided for the Validity Scales in Table 2, the Restructured Clinical Scales in Table 3, and the remaining substantive scales in Table 3.

MMPI-2

The MMPI-2 Clinical Scales were examined in this study because the vast majority of MMPI chronic pain research focuses on these scales (Tarescavage, 2015). The MMPI-2 Manual (Butcher et al., 2001) provides detailed information on the psychometrics of Clinical Scale scores in a variety of different samples. Of note, the Masculinity–Femininity and Introversion Clinical Scales were not investigated because they do not measure constructs relevant to psychopathology. Reliability estimates of MMPI-2 Clinical Scale scores in the current sample are provided in Table 3.

Table 2 MMPI-2-RF Validity Scale Scores (N = 230)

MMPI-2-RF scale name (abbreviation)	α	M	SD
Inconsistent responding			
Variable response inconsistency (VRIN-r)	.17	51.6	9.7
True response inconsistency (TRIN-r)	.12	57.4	6.4
Overreporting Scales			
Infrequent responses (F-r)	.70	69.2	15.5
Infrequent psychopathology responses (Fp-r)	.37	52.6	10.8
Infrequent somatic responses (Fs)	.58	64.8	18.2
Symptom validity (FBS-r)	.67	71.2	13.2
Response bias (RBS)	.68	66.4	15.8
Underreporting Scales			
Uncommon virtues (L-r)	.44	53.4	9.3
Adjustment validity (K-r)	.65	46.7	9.4

MMPI-2-RF Minnesota Multiphasic Personality Inventory-2-Restructured Form, *RC* restructured clinical, *M* mean, *SD* standard deviation, α Cronbach's alpha

Depression, Anxiety, Stress Scales

The Depression, Anxiety, Stress Scales (DASS) (Lovibond & Lovibond, 1996) is a 42-item self-report measure of mood problems. It has three scales measuring depression, anxiety, and generalized distress. Scores from the test have demonstrated adequate internal consistency reliability in a variety of settings (Lovibond & Lovibond, 1996) and have documented sensitivity to change (Page, Hooke, & Morrison, 2007).

Pain Variables

Patients rated the severity of their pain at intake on 11-point scales (0–10). The average pain intensity was 6.7 ($SD = 2.0$). They also reported the duration of their pain, which ranged from 1 to 63 years. The average duration of pain was 11.6 years ($SD = 9.9$).

Procedure

Upon admission to the Chronic Pain Rehabilitation Program (CPRP), all patients were given the MMPI-2 as part of a battery of tests including self-reported measures, and extensive patient and collateral interviews. The MMPI-2 was used, in part, to render psychological diagnoses, but largely to guide treatment. Participants were not excluded from treatment based on MMPI-2 scores. Evaluation dates ranged from 1999 to 2008. The CPRP is a comprehensive, intensive, interdisciplinary program that includes physical therapy, occupational therapy, group and individual psychological therapy, and medication management, including the weaning of all addicting substances including opioids and

Table 3 MMPI-2-RF RC Scales and MMPI-2 Clinical Scales descriptives and correlations with latent emotional distress outcome controlling for intake emotional distress, gender, age, pain intensity, and pain duration

MMPI-2-RF Scale name (abbreviation)	α	M	SD	$r_{outcome}$	MMPI-2 Scale name (abbreviation)	α	M	SD	$r_{outcome}$	d
Restructured Clinical Scales					Clinical Scales					
Demoralization (RCd)	.89	62.4	11.0	.29***	–	–	–	–	–	–
Somatic complaints (RC1)	.80	71.1	11.4	.13	Hypochondriasis (CS1)	.80	76.5	10.3	-.05	.50
Low positive emotions (RC2)	.70	62.3	11.5	.13	Depression (CS2)	.41	76.6	13.3	.04	1.15
Cynicism (RC3)	.83	48.2	10.4	.34***	Hysteria (CS3)	.65	78.7	13.8	-.10	2.52 ^a
Antisocial behavior (RC4)	.73	51.5	10.0	.21**	Psychopathic deviate (CS4)	.64	62.8	11.9	.02	1.03
Ideas of persecution (RC6)	.66	53.5	11.2	.26***	Paranoia (CS6)	.43	60.8	11.8	.13	.63
Dysfunctional negative emotions (RC7)	.85	54.1	11.6	.38***	Psychasthenia (CS7)	.88	67.9	12.0	.15	1.17
Aberrant experiences (RC8)	.67	53.8	10.2	.31***	Schizophrenia (CS8)	.88	67.4	12.0	.20*	1.23
Hypomanic activation (RC9)	.77	46.0	9.1	.30***	Mania (CS9)	.49	52.0	9.3	.04	.65

Consistent with interpretive guidelines, Clinical Scales are K-corrected (Graham, 2012)

Italicized findings attained a statistically significant and clinically meaningful effect size ($r \geq .20$)

MMPI-2-RF Minnesota Multiphasic Personality Inventory-2-Restructured Form, RC restructured clinical, M mean, SD standard deviation, $r_{outcome}$ correlation with outcome variable, d Cohen's d comparing mean RC versus CS scores, α Cronbach's alpha

^aRC3 measures one component of CS3, naivete, which is reversed to measure cynicism. Therefore, the RC3/CS3 comparisons reflect not only differences in psychometric properties but also in direction of keying

* $p < .05$; ** $p < .01$; *** $p < .001$

benzodiazepines as well as sedative hypnotics. Education about addiction and chemical dependency was offered as needed. The average length of stay is 3½ weeks and the treatment day extends from 7:30 am–5 pm, 5 days a week. Use of the sample was approved by an institutional review board.

Analysis Plan

MMPI-2 and MMPI-2-RF Mean Score Comparisons

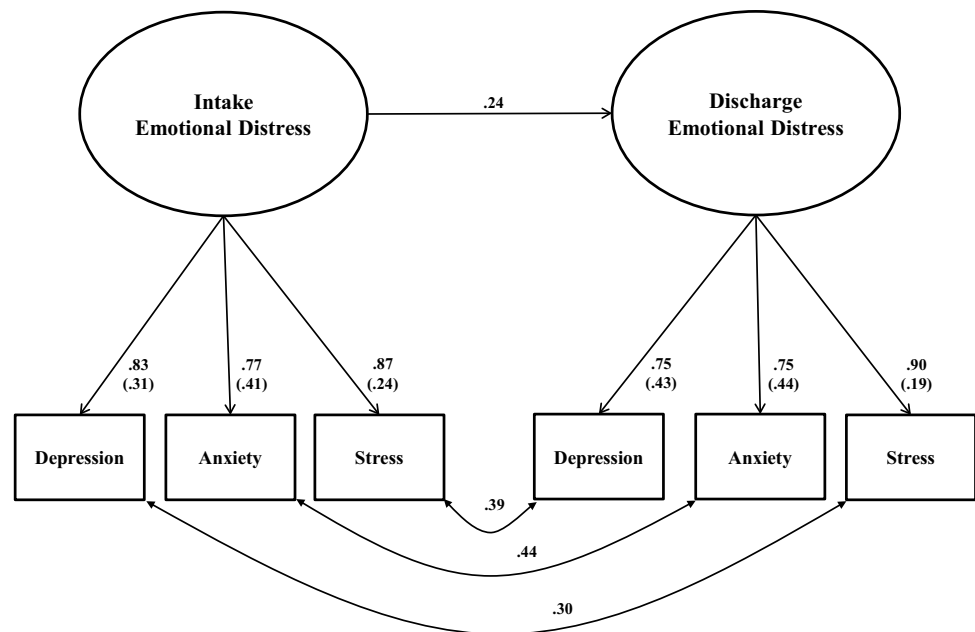
We first examined mean score differences between the MMPI-2 Clinical Scales and MMPI-2-RF RC Scales in the sample. We compared the values using Cohen's d , with values of .30, .50, and .80 representing small, medium, and large differences, respectively (Cohen, 1992). This analysis was intended to build on research by McCord and Drerup (2011) who found that the structural problems of the MMPI-2 Clinical Scales limited their interpretive utility relative to the MMPI-2-RF RC Scales. These authors categorized a sample of chronic pain patients into depressed and nondepressed diagnostic groups. In the depressed group in their study, mean clinical elevations were observed for the following Clinical Scales: 1 (Hypochondriasis), 2 (Depression), 3 (Hysteria), 4 (Psychopathic Deviate), 6 (Paranoia), 7 (Psychasthenia), and 8 (Schizophrenia). In stark contrast to the Clinical Scale findings, mean RC scale elevations were observed in the depressed group for only RCd (demoralization), RC1 (somatic complaints), and RC2 (low positive emotions), demonstrating substantially improved discriminant validity and interpretive utility.

MMPI-2 and MMPI-2 RF Associations with Outcome

Associations between the MMPI-2 Clinical Scales and MMPI-2-RF substantive scales, on the one hand, and treatment outcomes, on the other, were examined next in a structural equation modeling framework. All the analyses were completed in *Mplus, Version 6.11* (Muthén & Muthén, 2010) using the Maximum Likelihood (ML) estimator. Model fit was evaluated using the χ^2 test, the Comparative Fit Index (CFI) (Bentler, 1990), the Tucker Lewis Index (TLI) (Bentler & Bonett, 1980), and the Root Mean Square of Approximation (RMSEA) (Steiger, 1998). Nonsignificant χ^2 tests ($p > .05$), CFI and TLI values greater than .90, and RMSEA values less than .08 are indicative of adequate fit (Bentler, 1990; Browne, Cudeck, Bollen, & Long, 1993; Vandenberg & Lance, 2000; Yu, 2002). However, because the χ^2 test is overpowered in larger samples like the current one, statistically significant findings do not necessarily indicate poor fit.

We first identified a measurement model for the outcome variable, with scales from the DASS being used to model latent emotional distress factors at intake and discharge. We specified the intake factor as a predictor of the discharge factor to control for baseline emotional functioning (see Fig. 1 for final parameter estimates), a method recommended by Little (2013). The indicators were approximately normally distributed, supporting use of maximum likelihood estimation. The resulting model fit adequately. Specifically, the χ^2 test was nonsignificant ($\chi^2[5] = 8.23, p = .14$), and the other fit indices were also adequate (CFI = .99; TLI = .99; RMSEA = .05, 95% CI = .00 to .087). As reported later, we next specified age, gender, pain intensity, and duration of

Fig. 1 Emotional distress measurement model. All parameter estimates are statistically significant ($p < .013$). Manifest variables are scales from the depression anxiety stress scales. Residuals are in parentheses



pain as predictors of discharge emotional distress to control for these variables, a method also recommended by Little (2013). To identify associations between the MMPI-2-RF scales and future outcome, we correlated the test's scales with the outcome variable at discharge. Traditional MMPI guidelines indicate that a correlation of .20 or greater is clinically meaningful (Graham, Ben-Porath, & McNulty, 1999). Consistent with this guideline, we only interpreted statistically significant correlations ($p < .05$) yielding a magnitude of .20 or greater.

MMPI-2 and MMPI-2 RF Associations with Recovery

Finally, we examined associations between MMPI-2 and MMPI-2-RF scores and emotional distress outcome in the context of the clinically significant change model of outcome measure progress developed by Jacobson and Truax (Jacobson & Truax, 1991). According to this model and in the context of this study, clinically significant recovery occurs when a distressed individual achieves a statistically significant and reliable change (i.e., 1.96 multiplied by the outcome measure's standard error of difference) that is more characteristic of the general (nondistressed) population than the patient (distressed) population. For the purposes of the current study, individuals were deemed to have recovered if their DASS total score decreased/improved by 9.03 points (Reliable Change Index) to a score below 30.38 (midpoint between nondistressed and distressed population).

The just mentioned values were derived from DASS general population normative data provided by Crawford and Henry (Crawford & Henry, 2003) as well as the intake DASS scores in the current study using formula provided

by Jacobson and Truax (1991). Specifically, for the Reliable Change Index, the general population normative data indicated a standard error of measurement of 3.26. Applying the following formula (Jacobson & Truax, 1991) yielded a value of 9.03 for the Reliable Change Index: $1.96 \times \sqrt{(2 \times 3.26^2)}$. For the midpoint between the non-distressed versus distressed samples, we utilized means and standard deviations from the general population normative sample of DASS scores ($M = 18.38$, $SD = 18.82$) and from the study sample of intake DASS Scores ($M = 48.41$, $SD = 28.28$). The following formula (Jacobson & Truax, 1991) yielded a value of 30.38 for the midpoint of the non-distressed population versus distressed population: $\{[(28.28 \times 18.38) + (18.82 \times 48.41)] / (18.82 + 28.28)\}$.

DASS Total Scores were available at both the intake and discharge time points for 181 members of the current sample. Overall, 91 of these individuals (50.3%) met the criterion for recovered, such that their DASS Total Score decreased by more than 9.03 points (Reliable Change Index) to a value less than 30.38 (nondistressed vs. distressed population midpoint). Of the 90 remaining individuals, 10 individuals (5.5%) had a change that was less than the Reliable Change Index, 6 individuals (3.1%) had a reliable change in the direction of deterioration, and 17 individuals (8.9%) had a reliable change in the direction of improvement but ultimately did not have a score below the nondistressed/distressed population midpoint of 30.38. These 33 individuals (18.2%) were therefore considered to have not recovered. The remaining 57 individuals (29.8%) had DASS intake and discharge scores that were both in the nondistressed range; therefore, they neither met criteria for recovered nor nonrecovered. We compared mean

MMPI-2 and MMPI-2-RF T-scores across the recovered ($n = 91$) and nonrecovered ($n = 33$) groups using t-tests and examined effect size using Cohen's d . Consistent with traditional MMPI guidelines (Graham, Ben-Porath, & McNulty, 1999), we interpreted comparisons that yielded a statistically significant ($p < .05$) and clinically meaningful effect size ($d \geq .40$, which is equivalent to $r \geq .20$).

Results

MMPI-2 and MMPI-2-RF Mean Score Comparisons

We present in Table 3 MMPI-2-RF RC Scale T-score means and standard deviations for this sample alongside MMPI-2 Clinical Scale T-score means and standard deviations. The sample produced clinically significant mean elevations (i.e., a score $\geq 65T$) on only RC1 (Somatic Complaints). The sample scored at or above 60T (one standard deviation above the general population mean) on RCd (Demoralization) and RC2 (Low Positive Emotions). In contrast, the sample produced clinically significant mean elevations on most of the MMPI-2 Clinical Scales, including CS1 (Hypochondriasis), CS2 (Depression), CS3 (Hysteria), CS7 (Psychasthenia), and CS8 (Schizophrenia). The CS4 (Psychopathic Deviate) and CS6 (Paranoia) scales approached the threshold of a clinical elevation (both T-score were greater than 60). Only CS9 (Mania) was within one standard deviation of the general population mean (i.e., a T-score less than 60T). In general, MMPI-2 Clinical Scale T-scores were substantially higher than RC Scale T-scores, with Cohen's d values ranging from .50 (RC1/CS1) to 2.52 (RC3/CS3). Of note, RC3 measures one component of CS3, naivete, which is reversed to measure cynicism. For the interested reader, descriptive statistics for the remaining MMPI-2-RF scales examined in this study are presented in Table 2.

Effects of Age, Gender, Pain Intensity, and Duration of Pain on Outcomes

As detailed in the analysis plan, we used age, gender, pain intensity, and duration of pain as predictors of emotional distress outcomes in order to control for these variables. In the emotional distress model, age was a significant predictor of outcome (Standardized coefficient = $-.24$, $p = .004$), but the remaining predictors were nonsignificant, including gender (Standardized coefficient = $-.05$, $p = .56$), pain intensity (Standardized coefficient = $-.13$, $p = .11$), and duration of pain (Standardized coefficient = $.04$, $p = .61$).

MMPI-2 and MMPI-2 RF Associations with Outcome

We present in Table 3 correlations between the MMPI-2-RF RC scales and MMPI-2 Clinical Scales and emotional distress treatment outcomes after controlling for baseline functioning in these areas, as well as age, gender, pain intensity, and pain duration. Most of the RC Scales were significantly, meaningfully associated with the criterion, with the following scales yielding correlates greater than .20: RCd (Demoralization), RC3 (Cynicism), RC4 (Antisocial Behavior), RC6 (Persecutory Ideation), RC7 (Dysfunctional Negative Emotions), RC8 (Aberrant Experiences), and RC9 (Hypomanic Activation). In contrast, only Clinical Scale 8 (Schizophrenia) was statistically and meaningfully associated with outcome. Regarding the rest of the MMPI-2-RF substantive scales (see Table 4), all five substantive domains were represented as predictors of poor emotional distress outcomes, including emotional dysfunction (EID, HLP, NFC, STW, AXY, BRF, and NEGE-r), behavioral dysfunction (BXD, AGG, ACT, and DISC-r), thought dysfunction (THD and PSYC-r), interpersonal dysfunction (SHY and DSF), and somatic/cognitive dysfunction (MLS and COG).

MMPI-2 and MMPI-2 RF Associations with Recovery

Finally, we present in Table 5 mean score differences for the MMPI-2-RF RC Scales and MMPI-2 Clinical Scales across groups of individuals who either recovered following treatment according to the DASS ($n = 91$) or those who did not recover according to this measure ($n = 33$). Most of the RC Scales were significantly and meaningfully different across the groups, with the following scales yielding Cohen's d effect sizes greater than .40 (indicating higher-pretreatment scores in the nonrecovered group): RCd (Demoralization), RC3 (Cynicism), RC4 (Antisocial Behavior), RC6 (Persecutory Ideation), RC7 (Dysfunctional Negative Emotions), RC8 (Aberrant Experiences), and RC9 (Hypomanic Activation). In contrast, only Clinical Scale 9 (Mania) was statistically and meaningfully higher in the nonrecovered group of the two groups. Regarding the rest of the MMPI-2-RF substantive scales (see Table 5), most of the substantive domains were significantly and meaningfully higher in the nonrecovered group, including behavioral dysfunction (JCP, SUB, AGG, DISC-r), thought dysfunction (PSYC-r), interpersonal dysfunction (DSF), and somatic/cognitive dysfunction (COG). Overall and in general terms, mean T-scores in the nonrecovered group approximated a clinical elevation (65T) on scales measuring emotional dysfunction and somatic/cognitive complaints, whereas they typically approximated a score of 55T for scales measuring behavioral dysfunction in this same group. Thought dysfunction scores in the nonrecovered group approximated a T-score of 60.

Table 4 MMPI-2-RF higher-order, specific problems, and personality psychopathology-5 scale descriptives and correlations with latent emotional distress outcome controlling for intake emotional distress, gender, age, pain intensity, and pain duration

MMPI-2-RF Scale name (abbreviation)	α	M	SD	$r_{outcome}$
Higher-Order Scales				
Emotional/internalizing dysfunction (EID)	.90	62.2	11.6	.30***
Thought dysfunction (THD)	.64	53.0	9.9	.25***
Behavioral/externalizing dysfunction (BXD)	.77	48.6	9.8	.27***
Specific Problems Scales				
Malaise (MLS)	.51	76.4	7.9	.21**
Gastrointestinal complaints (GIC)	.71	64.8	16.6	-.07
Head pain complaints (HPC)	.74	64.5	10.3	.10
Neurological complaints (NUC)	.66	69.9	13.2	.15
Cognitive complaints (COG)	.77	63.2	14.1	.39***
Suicidal/death ideation (SUI)	.67	54.8	16.0	.04
Helplessness/hopelessness (HLP)	.60	52.7	12.4	.31***
Self-doubt (SFD)	.75	58.7	12.0	.05
Inefficacy (NFC)	.76	53.9	11.6	.29***
Stress/worry (STW)	.60	56.7	11.3	.30***
Anxiety (AXY)	.62	57.7	14.8	.26**
Anger proneness (ANP)	.76	53.5	12.0	.20*
Behavior-restricting fears (BRF)	.52	53.9	12.0	.25***
Multiple specific fears (MSF)	.70	50.8	9.4	.09
Juvenile conduct problems (JCP)	.66	50.3	10.9	.19*
Substance abuse (SUB)	.63	50.3	10.1	.15
Aggression (AGG)	.61	48.6	9.4	.22**
Activation (ACT)	.59	47.9	9.9	.26***
Family problems (FML)	.68	53.0	11.1	.17*
Interpersonal passivity (IPP)	.70	52.2	10.7	.01
Social avoidance (SAV)	.80	53.0	11.0	.16*
Shyness (SHY)	.77	51.2	10.8	.26***
Disaffiliativeness (DSF)	.60	52.9	12.7	.29***
Personality-Psychopathology-5 Scales				
Aggressiveness-revised (AGGR-r)	.76	47.3	9.6	.03
Psychoticism-revised (PSYC-r)	.62	52.9	9.8	.31***
Disconstraint-revised (DISC-r)	.75	48.2	9.8	.25**
Negative emotionality/neuroticism-revised (NEGE-r)	.81	57.9	12.3	.36***
Introversion/low positive emotionality-revised (INTR-r)	.79	57.3	12.2	.05

* $p < .05$; ** $p < .01$; *** $p < .001$

Italicized findings attained a statistically significant and clinically meaningful effect size ($r \geq .20$)

MMPI-2-RF Minnesota Multiphasic Personality Inventory-2-Restructured Form, *RC* restructured clinical, *M* mean, *SD* standard deviation, $r_{outcome}$ correlation with outcome variable, *d* Cohen's *d* comparing mean RC versus CS scores, α Cronbach's alpha

Discussion

The purpose of this study was to examine the relative utility of the MMPI-2-RF substantive scales compared to the MMPI-2 Clinical Scales in the prediction of emotional distress outcomes among patients with chronic low back pain undergoing intensive outpatient treatment. Descriptive analyses of MMPI-2-RF scores indicated that the current sample reported relatively high levels of somatic problems, as well as mood disorder-related symptomatology.

However, in comparisons of the RC and MMPI-2 Clinical Scales, scores on the latter suggested substantially more severity and variability in psychopathology. Finally, scales from all domains from the MMPI-2-RF demonstrated associations with psychological distress outcome, whereas MMPI-2 Clinical Scale scores generally did not demonstrate meaningful associations. These results were consistent with clinically relevant comparisons across groups of individuals who at discharge either recovered or did not recover from their intake level of psychological

Table 5 Mean comparisons between emotional distress recovered and not recovered groups

Scale name (abbreviation)	Recovered (<i>n</i> =91)		Not recovered (<i>n</i> =33)		Statistical comparisons		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	Sig.	<i>d</i>
MMPI-2-RF Higher-Order Scales							
Emotional/internalizing dysfunction (EID)	64.1	11.1	67.8	11.0	1.648	.102	.34
Thought dysfunction (THD)	53.6	10.2	57.1	9.4	1.721	.088	.36
Behavioral/externalizing dysfunction (BXD)	46.5	8.1	56.7	12.0	5.424	<.001	1.02
MMPI-2-RF restructured Clinical Scales							
Demoralization (RCd)	64.2	10.1	68.8	9.3	2.269	.025	.47
Somatic complaints (RC1)	72.5	11.1	71.7	11.5	−0.342	.733	−.07
Low positive emotions (RC2)	63.0	11.0	63.9	13.4	0.386	.701	.07
Cynicism (RC3)	48.0	9.4	56.5	14.0	3.871	<.001	.73
Antisocial behavior (RC4)	49.6	9.3	57.5	11.5	3.914	<.001	.76
Ideas of persecution (RC6)	53.8	11.7	59.9	11.6	2.596	.011	.53
Dysfunctional negative emotions (RC7)	55.6	10.6	61.3	12.0	2.533	.013	.50
Aberrant experiences (RC8)	54.3	10.0	59.0	10.1	2.312	.022	.47
Hypomanic activation (RC9)	45.9	8.8	53.1	11.6	3.652	<.001	.70
MMPI-2 Clinical Scales							
Hypochondriasis (CS1)	77.3	10.2	75.5	12.6	−0.804	.423	−.16
Depression (CS2)	79.5	12.8	76.5	13.1	−1.152	.252	−.23
Hysteria (CS3)	80.4	14.8	76.8	15.7	−1.181	.240	−.24
Psychopathic deviate (CS4)	63.1	12.4	64.9	11.1	0.726	.469	.15
Paranoia (CS6)	62.5	12.4	63.4	11.7	0.362	.718	.07
Psychasthenia (CS7)	69.9	10.9	72.2	11.4	1.036	.302	.21
Schizophrenia (CS8)	68.2	11.8	72.7	10.9	1.931	.056	.40
Mania (CS9)	51.1	8.6	55.4	9.6	2.410	.017	.48
MMPI-2-RF specific problems Scales							
Malaise (MLS)	76.8	7.6	79.0	7.6	1.447	.150	.29
Gastrointestinal complaints (GIC)	66.4	16.5	64.4	16.3	−0.616	.539	−.13
Head pain complaints (HPC)	65.5	10.3	64.5	9.1	−0.502	.617	−.11
Neurological complaints (NUC)	71.0	12.9	71.3	12.6	0.138	.890	.03
Cognitive complaints (COG)	63.6	13.0	70.9	12.2	2.820	.006	.58
Suicidal/death ideation (SUI)	56.4	17.4	56.9	16.3	0.149	.882	.03
Helplessness/hopelessness (HLP)	52.6	11.3	57.5	14.7	1.956	.053	.38
Self-doubt (SFD)	60.5	11.8	61.1	11.8	0.248	.805	.05
Inefficacy (NFC)	55.4	11.2	58.2	10.1	1.252	.213	.26
Stress/worry (STW)	58.2	10.0	60.7	11.9	1.183	.239	.23
Anxiety (AXY)	58.7	14.2	64.8	17.4	1.971	.051	.38
Anger proneness (ANP)	55.2	11.7	59.3	13.2	1.660	.100	.33
Behavior-restricting fears (BRF)	54.9	12.1	58.3	12.6	1.377	.171	.28
Multiple specific fears (MSF)	51.6	10.4	49.6	7.0	−1.045	.298	−.24
Juvenile conduct problems (JCP)	49.0	10.6	56.4	12.2	3.307	.001	.65
Substance abuse (SUB)	47.8	8.0	56.9	12.6	4.782	<.001	.89
Aggression (AGG)	48.4	8.5	54.8	12.9	3.188	.002	.60
Activation (ACT)	49.0	9.8	50.9	10.9	0.936	.351	.19
Family problems (FML)	53.5	12.1	55.7	12.0	0.873	.384	.18
Interpersonal passivity (IPP)	50.5	9.3	52.3	11.6	0.871	.385	.17
Social avoidance (SAV)	52.9	11.1	56.3	12.2	1.434	.154	.28
Shyness (SHY)	51.8	10.9	55.6	11.7	1.692	.093	.34
Disaffiliativeness (DSF)	53.4	12.7	60.2	17.2	2.391	.018	.46

Table 5 (continued)

Scale name (abbreviation)	Recovered (<i>n</i> =91)		Not recovered (<i>n</i> =33)		Statistical comparisons		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	Sig.	<i>d</i>
MMPI-2-RF personality-psychopathology-5 Scales							
Aggressiveness-revised (AGGR-r)	48.1	9.5	49.8	11.9	0.833	.406	.16
Psychoticism-revised (PSYC-r)	53.3	10.2	57.6	8.7	2.147	.034	.45
Disconstraint-revised (DISC-r)	45.4	7.4	56.1	11.7	6.014	<.001	1.12
Negative emotionality/neuroticism-revised (NEGE-r)	59.9	11.8	64.1	11.9	1.734	.085	.35
Introversion/low positive emotionality-revised (INTR-r)	58.0	12.9	57.6	13.1	-0.154	.878	-.03

Consistent with interpretive guidelines, Clinical Scales are K-corrected (Graham, 2012)

Italicized findings attained a statistically significant and clinically meaningful effect size ($d \geq .40$)

MMPI-2-RF Minnesota Multiphasic Personality Inventory-2-Restructured Form, *M* mean, *SD* standard deviation, *d* Cohen's *d*

distress. Several aspects of these findings warrant further discussion.

As just noted, mean comparisons of the RC and Clinical Scales tended to demonstrate much higher scores across the Clinical Scales. Moreover, the pattern of mean RC Scale scores appeared more consistent with the types of psychopathology present in the sample. For example, approximately half of the sample had a comorbid major depressive disorder, which is consistent with the observed mean subthreshold elevations of approximately 62T on RCd (demoralization) and RC2 (low positive emotions). One would expect overall sample mean scores of 65T or higher only if the entire sample had major depressive disorder, which is not the case. Along the same lines, the observed mean elevation on RC1 (somatic complaints) is to be expected among a sample of patients with chronic low back pain. In contrast, the Clinical Scales demonstrated a substantially higher mean score on its measure of depression (CS2); indeed, it was consistent with the 99th percentile in the general population. Moreover, the MMPI-2 Clinical Scales evidenced subthreshold to threshold mean elevations on scales that did not reflect psychopathology commonly found in this population (Psychopathic Deviate, Paranoia, and Schizophrenia). Taken together and in line with past research (McCord & Drerup, 2011), the MMPI-2-RF RC Scales appeared to evidence more interpretive utility and particularly discriminant validity than did the MMPI-2 Clinical Scales in this setting.

Correlation comparisons across the MMPI-2-RF and MMPI-2 Scales more directly demonstrated the limitations of the Clinical Scales. Whereas 7 of the 9 RC Scales (and several other MMPI-2-RF Scales) were significantly, meaningfully associated with the outcome variable, only one MMPI-2 Clinical Scale (Schizophrenia) demonstrated a significant association. The latter finding is consistent with most past research indicating that MMPI/MMPI-2 Scale scores are not associated with outcomes in this setting

(McGill, Lawlis, Selby, Mooney, & McCoy, 1983; Moore, Armentrout, Parker, & Kivlahan, 1986). However, notwithstanding the limitations of its predecessors, the newest version of the MMPI—the MMPI-2-RF—was meaningfully associated with outcomes in this study, likely owing to its improved psychometrics and convergence with modern models of psychopathology (Sellbom, Ben-Porath, & Bagby, 2008).

Findings from mean comparisons of recovered and not recovered groups using a clinically significant change framework generally converged with the correlational results. However, these findings have more direct clinical implications, as robust correlation coefficients do not necessarily translate to actual treatment gains or losses as a result of higher scores on a Scale. Moreover, the mean comparisons across recovered and not recovered groups enable practitioners to identify which scores on a scale are most likely to be associated with a problematic outcome. Accordingly, the results indicated that practitioners should interpret the MMPI-2-RF with some flexibility rather than strictly adhere to interpretations of 65T or higher (for the purposes of identifying individuals at risk for poor emotional distress outcomes). For example, BXD demonstrated a large effect size (Cohen's $d = 1.02$) in differentiating recovered and not recovered groups, with a mean score of 57T in the nonrecovered group.

Treatment Implications

The findings of this study can be used to assist with interpretation of the MMPI-2-RF in this setting, such that individuals with marked scores on scales associated with poor outcome can be provided targeted interventions to increase their chances of success. In order to inform treatment, a brief description of how scales with the most robust findings can impede treatment are described next, using the test's

interpretive manual as a guide (Ben-Porath & Tellegen, 2008/2011).

Although individuals presenting with higher scores on the Emotional/internalizing scales (Emotional/internalizing dysfunction, dysfunctional negative emotions, helplessness/hopelessness, stress/worry, and negative emotionality/neuroticism-revised) might be initially motivated to engage in treatment due to their distress, they may disengage after their distress levels begin to subside. Scores on scales that assess negatively emotionality and its facets had the strongest associations with poorer emotional distress functioning post-treatment, indicating the importance of targeting these constructs in treatment planning. These findings converge with those of Marek, Block, and Ben-Porath (2014), who found that MMPI-2-RF markers of emotional distress, depression, and negative emotionality were associated with pain disability outcomes after spinal cord surgery.

Individuals with high scores on the Behavioral/externalizing dysfunction, hypomanic activation, and disconstraint scales are likely to be noncompliant with treatment efforts due to excessive activation, antisocial orientation, or impulsivity. Unusual thoughts and cognitions, as evidenced by moderately elevated scores on scales such as Aberrant Experiences and psychoticism-revised, may interfere with treatment as well. Individuals with greater cynical and disaffiliative attitudes may have difficulty forming a therapeutic relationship, and they may have less access to socially supportive others, which may account for the associations between RC3 and DSF with poor outcomes. The RC3 findings are consistent with past research showing high scores on this scale can lead poor treatment engagement among national guard soldiers (Arbisi, Polusny, Erbes, Thuras, & Reddy, 2011). Finally, patients with higher scores on cognitive complaints are likely to have a low tolerance for frustration, which could impede the benefits of participating in an intensive chronic pain treatment program.

Although not a focus of the current study, it is worth mentioning that the MMPI-2-RF Validity Scales can also aid in treatment efforts by assessing a patient's response style. In the current study, approximately 13% of patients from the initial sample evidenced an invalid response style on the MMPI-2-RF. That is, these individuals demonstrated inconsistent responding or overreporting of psychopathology or somatic/cognitive complaints. Past research in other settings has demonstrated that MMPI-2-RF identified overreporting is likely to generalize to other aspects of the assessment (Forbey, Lee, Ben-Porath, Arbisi, & Gartland, 2013). Thus, in situations where an individual produces an invalid protocol, practitioners should be aware of the limitations of concurrently obtained patient data. Invalid protocols may also have implications for treatment, as demonstrated by studies in mental health settings (Anestis, Finn, Gottfried, Arbisi, & Joiner, 2014). However, future research on the utility of

the MMPI-2-RF Validity Scales for this purpose in chronic pain treatment is needed.

Limitations and Conclusion

Limitations of the current study provide direction for future research. The sample was entirely composed of low back pain patients. Although low back pain is the most common complaint of individuals in this setting, it would be useful to see whether these findings generalize to other diagnoses, such as fibromyalgia, chronic fatigue syndrome, migraine headaches, multiple sclerosis, and other neurological disorders. Moreover, replication of these results in other types of chronic pain treatment settings is necessary, as the patients in the current study participated in a structured interdisciplinary pain treatment program that was very intensive. Although such a program is recommended, it does not represent the treatment received by most pain patients. Additionally, given the intensive nature of the program, the severity of chronic pain and comorbid diagnoses such as prescription opioid abuse may be greater than what is typically observed in outpatient settings. The results of the study were analyzed in the context of exploration (Reichenbach, 1938), given the limited available research on the MMPI-2-RF and chronic pain treatment outcomes. Consequently, replication is needed. Moreover, the current study utilized only one outcome measure—future research with multiple and varied methods of outcomes assessment are indicated. Along the same lines, some patient characteristics were not assessed but could be relevant to the generalizability of the study (e.g., treatment satisfaction). Finally, investigating associations of MMPI-2-RF scores with other indicators of outcome would be beneficial and may provide additional insight into the utility of the instrument in assessments of patients with chronic low back pain. For example, the Externalizing Scales may be useful in the prediction of treatment nonadherence.

These limitations notwithstanding, the current study is the first to investigate the utility of the MMPI-2-RF in the prediction of treatment outcomes among chronic pain patients undergoing conservative treatments. Overall, the results of this study provide preliminary support for the use of the MMPI-2-RF among patients with chronic low back pain and should ease long-standing concerns that an MMPI instrument is not useful in this setting.

Compliance with Ethical Standards

Conflict of interest Yossef Ben-Porath is a paid consultant to the MMPI-2-RF publisher, the University of Minnesota Press, and a distributor for Pearson Assessments. He receives royalties on sales of MMPI-2-RF materials and research Grants from the MMPI-2-RF publisher. Anthony Tarescavage also receives research funding from the MMPI-2-RF publisher. The authors thank Dr. Kelly Huffman for her help with

institutional review board processing and approval. Judith Scheman declares that she does not have any conflict of interest.

Human and Animal Rights All procedures were in accordance with the ethical standards of the institutional research committees and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. Washington, DC: American Psychiatric Publisher.
- Anestis, J. C., Finn, J. A., Gottfried, E., Arbisi, P. A., & Joiner, T. E. (2014). Reading the road signs the utility of the MMPI-2 Restructured Form validity scales in prediction of premature termination. *Assessment, 22*, 279–288.
- Arbisi, P. A., Polusny, M. A., Erbes, C. R., Thuras, P., & Reddy, M. K. (2011). The Minnesota Multiphasic Personality Inventory-2 Restructured Form in National Guard soldiers screening positive for posttraumatic stress disorder and mild traumatic brain injury. *Psychological Assessment, 23*(1), 203. <https://doi.org/10.1037/a0021339>.
- Ben-Porath, Y. S., & Tellegen, A. (2008/2011). *MMPI-2-RF: Manual for administration, scoring and interpretation*. Minneapolis, MN: University of Minnesota Press.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238–246.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588–606. <https://doi.org/10.1037/0033-2909.88.3.588>.
- Bradley, L. A. (1995). Biopsychosocial model and the Minnesota Multiphasic Personality Inventory: Even psychologists make cognitive errors. *Pain Forum, 4*(2), 97–100. [https://doi.org/10.1016/S1082-3174\(11\)80006-1](https://doi.org/10.1016/S1082-3174(11)80006-1).
- Browne, M. W., Cudeck, R., Bollen, K. A., & Long, J. S. (1993). Alternative ways of assessing model fit. *Sage Focus Editions, 154*, 136–136.
- Butcher, J., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, G. W., & Kaemmer, B. (2001). *Minnesota Multiphasic Personality Inventory-2: Manual for administration, scoring, and interpretation*. Minneapolis, MN: University of Minnesota Press.
- Chou, R., Qaseem, A., Snow, V., Casey, D., Cross, J. T. Jr., Shekelle, P., & Owens, D. K. (2007). Diagnosis and treatment of low back pain: A joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Annals of Internal Medicine, 147*(7), 478–491.
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159.
- Crawford, J. R., & Henry, J. D. (2003). The Depression Anxiety Stress Scales (DASS): Normative data and latent structure in a large non-clinical sample. *British Journal of Clinical Psychology, 42*(2), 111–131.
- Forbey, J. D., Lee, T. T., Ben-Porath, Y. S., Arbisi, P. A., & Gartland, D. (2013). Associations between MMPI-2-RF validity scale scores and extra-test measures of personality and psychopathology. *Assessment, 20*, 448.
- Freburger, J. K., Holmes, G. M., Agans, R. P., Jackman, A. M., Darter, J. D., Wallace, A. S., ... Carey, T. S. (2009). The rising prevalence of chronic low back pain. *Archives of Internal Medicine, 169*(3), 251–258. <https://doi.org/10.1001/archinternmed.2008.543>.
- Gatchel, R. J., McGeary, D. D., McGeary, C. A., & Lippe, B. (2014). Interdisciplinary chronic pain management: Past, present, and future. *American Psychologist, 69*(2), 119.
- Gatchel, R. J., Peng, Y. B., Peters, M. L., Fuchs, P. N., & Turk, D. C. (2007). The biopsychosocial approach to chronic pain: Scientific advances and future directions. *Psychological Bulletin, 133*(4), 581–624. <https://doi.org/10.1037/0033-2909.133.4.581>.
- Graham, J. R. (2012). *The MMPI-2: Assessing personality and psychopathology*. New York: Oxford University Press.
- Graham, J. R., Ben-Porath, Y. S., & McNulty, J. L. (1999). *MMPI-2 correlates for outpatient community mental health settings*. Minneapolis, MN: University of Minnesota Press.
- Guo, H., Tanaka, S., Halperin, W. E., & Cameron, L. L. (1999). Back pain prevalence in US industry and estimates of lost workdays. *American Journal of Public Health, 89*(7), 1029–1035.
- Hathaway, S. R., & McKinley, J. C. (1943). *Minnesota Multiphasic Personality Inventory*. Minneapolis, MN: University of Minnesota Press.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19.
- Katz, J. N. (2006). Lumbar disc disorders and low-back pain: Socioeconomic factors and consequences. *The Journal of Bone & Joint Surgery, 88*(suppl_2), 21–24. <https://doi.org/10.2106/JBJS.E.01273>.
- Keefe, F. J., Lefebvre, J. C., & Beaupre, P. M. (1995). The Minnesota Multiphasic Personality Inventory in chronic pain: Security blanket or sound investment? *Pain Forum, 4*(2), 101–103. [https://doi.org/10.1016/S1082-3174\(11\)80007-3](https://doi.org/10.1016/S1082-3174(11)80007-3).
- Kessler, R. C., Berglund, P., Chiu, W. T., Demler, O., Heeringa, S., Hiripi, E., ... Zheng, H. (2004). The US National Comorbidity Survey Replication (NCS-R): Design and field procedures. *International Journal of Methods in Psychiatric Research, 13*(2), 69–92.
- LaBar, G. (1992). A battle plan for back injury prevention. *Occupational Hazards, 11*, 29–33.
- Little, T. D. (2013). *Longitudinal structural equation modeling*. New York: Guilford Press.
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the beck depression and anxiety inventories. *Behaviour Research and Therapy, 33*(3), 335–343.
- Lovibond, S. H., & Lovibond, P. F. (1996). *Manual for the depression anxiety stress scales*. Sydney: Psychology Foundation of Australia.
- Main, C. J., & Spanswick, C. C. (1995). Personality assessment and the Minnesota Multiphasic Personality Inventory: 50 years on: Do we still need our security blanket? *Pain Forum, 4*(2), 90–96. [https://doi.org/10.1016/S1082-3174\(11\)80005-X](https://doi.org/10.1016/S1082-3174(11)80005-X).
- Manchikanti, L., Cash, K. A., Damron, K. S., Manchukonda, R., Pampati, V., & McManus, C. D. (2006). Controlled substance abuse and illicit drug use in chronic pain patients: An evaluation of multiple variables. *Pain Physician, 9*(3), 215–225.
- Marek, R. J., Block, A. R., & Ben-Porath, Y. S. (2014). The Minnesota Multiphasic Personality Inventory-2-Restructured Form: (MMPI-2-RF): Incremental validity in predicting early postoperative outcomes in spine surgery candidate. *Psychological Assessment, 27*(1), 114–124. <https://doi.org/10.1037/PAS0000035>.
- McCord, D. M., & Drerup, L. C. (2011). Relative practical utility of the Minnesota Multiphasic Personality Inventory-2 Restructured Clinical Scales versus the Clinical Scales in a chronic pain patient sample. *Journal of Clinical and Experimental Neuropsychology, 33*(1), 140–146. <https://doi.org/10.1080/13803395.2010.495056>.

- McGill, J. C., Lawlis, G. F., Selby, D., Mooney, V., & McCoy, C. E. (1983). The relationship of Minnesota Multiphasic Personality Inventory (MMPI) profile clusters to pain behaviors. *Journal of Behavioral Medicine*, 6(1), 77–92. <https://doi.org/10.1007/BF00845278>.
- McNeil, J. M., & Binette, J. (2001). Prevalence of disabilities and associated health conditions among adults—United States, 1999. *Journal of the American Medical Association*, 285(12), 1571–1572.
- 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(1), 51–63. <https://doi.org/10.1007/BF00844644>.
- Muthén, L., & Muthén, B. (2010). *Mplus user's guide, v 6.1*. Los Angeles: Muthen and Muthen, UCLA.
- Page, A. C., Hooke, G. R., & Morrison, D. L. (2007). Psychometric properties of the Depression Anxiety Stress Scales (DASS) in depressed clinical samples. *British Journal of Clinical Psychology*, 46(3), 283–297.
- Piotrowski, C. (1998). Assessment of pain: A survey of practicing clinicians. *Perceptual and Motor Skills*, 86(1), 181–182.
- Piotrowski, C., & Lubin, B. (1990). Assessment practices of health psychologists: Survey of APA division 38 clinicians. *Professional Psychology: Research and Practice*, 21(2), 99–106. <https://doi.org/10.1037//0735-7028.21.2.99>.
- Reichenbach, H. (1938). *Experience and prediction: An analysis of the foundations and the structure of knowledge*. Chicago: University of Chicago Press.
- Rubin, D. I. (2007). Epidemiology and risk factors for spine pain. *Neurologic Clinics*, 25(2), 353–371. <https://doi.org/10.1016/j.ncl.2007.01.004>.
- Sellbom, M., Ben-Porath, Y. S., & Bagby, R. M. (2008). On the hierarchical structure of mood and anxiety disorders: Confirmatory evidence and elaboration of a model of temperament markers. *Journal of Abnormal Psychology*, 117(3), 576–590. <https://doi.org/10.1037/a0012536>.
- Steiger, J. H. (1998). A note on multiple sample extensions of the RMSEA fit index. *Structural Equation Modeling—a Multidisciplinary Journal*, 5(4), 411–419. <https://doi.org/10.1080/10705519809540115>. doi.
- Tarescavage, A. M. (2015). Predicting treatment outcomes among low back pain patients using the Minnesota Multiphasic Personality Inventory-2-Restructured Form. (Dissertation, Ph. D.), Kent State University, Kent, OH. Retrieved from http://rave.ohiolink.edu/etdc/view?acc_num=kent1447070671.
- Tellegen, A., & Ben-Porath, Y. S. (2008/2011). *Minnesota multiphasic personality inventory-2-Restructured Form: Technical manual*. Minneapolis, MN: University of Minnesota Press.
- Turk, D. C., & Fernandez, E. (1995). Personality assessment and the Minnesota Multiphasic Personality Inventory in chronic pain: Underdeveloped and overexposed. *Pain Forum*, 4(2), 104–107. [https://doi.org/10.1016/S1082-3174\(11\)80008-5](https://doi.org/10.1016/S1082-3174(11)80008-5).
- Van der Heijden, P. T., Egger, J. I., & Derksen, J. J. (2010). Comparability of scores on the MMPI-2-RF scales generated with the MMPI-2 and MMPI-2-RF booklets. *Journal of Personality Assessment*, 92(3), 254–259. <https://doi.org/10.1080/00223891003670208>.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4–70.
- Von Korff, M., Crane, P., Lane, M., Miglioretti, D. L., Simon, G., Saunders, K., ... Kessler, R. (2005). Chronic spinal pain and physical-mental comorbidity in the United States: Results from the national comorbidity survey replication. *PAIN*, 113(3), 331–339. <https://doi.org/10.1016/j.pain.2004.11.010>.
- Yu, C. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes*. Los Angeles: University of California Los Angeles.