Psychological Assessment
2005, Vol. 17, No. 3, 345–358

1040-3590/05/$12.00 DOI: 10.1037/1040-3590.17.3.345

Psychometric Evaluation of the Restructured Clinical Scales of the MMPI–2

Leonard J. Simms
University at Buffalo, The State University of New York

Alex Casillas, Lee Anna Clark, and David Watson
University of Iowa

Bradley N. Doebbeling
Richard L. Roudebush Veterans' Affairs Medical Center, Regenstrief Institute for Health Care, and Indiana University School of Medicine

Conceptual overlap and heterogeneity have long been noted as weaknesses of the Minnesota Multiphasic Personality Inventory's clinical scales. Restructured clinical (RC) scales recently were developed to address these concerns (A. Tellegen et al., 2003). The authors evaluated the psychometric properties of the RC scales in psychology clinic clients ($N = 285$) and military veterans ($N = 567$). The RC scales were as internally consistent as the clinical scales and correlated strongly with their original counterparts (except for RC3/Hysteria). They also were less intercorrelated, produced conceptually clearer relations with measures of personality and psychopathology, and yielded somewhat greater incremental utility than the clinical scales. Thus, the RC scales demonstrated several psychometric strengths while utilizing 60% fewer items, but the 2 sets of scales cannot be used interchangeably. Interpretive considerations are discussed.

Keywords: Minnesota Multiphasic Personality Inventory—2, restructured clinical scales, incremental validity, personality assessment

Conceptual overlap and heterogeneity have long been noted as weaknesses of the clinical scales of the Minnesota Multiphasic Personality Inventory (MMPI; e.g., Helmes & Reddon, 1993; Horn, Wanberg, & Appel, 1973; Waller, 1999), both in the original MMPI (Hathaway & McKinley, 1943) and in the MMPI—2nd Edition (MMPI–2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989; Butcher et al., 2001). Because maintaining continuity with the large fund of empirical research completed on the original MMPI was given a high priority, the revised MMPI–2 did little to correct these problems. However, Tellegen et al. (2003) attempted to address these concerns by developing a set of restructured clinical (RC) scales. The manual for the RC scales includes basic psychometric data derived from the normative sample as well as several psychiatric settings, but no independent evaluations of these new scales have yet appeared in the literature. Clearly, additional data are needed to evaluate these scales before clinicians and researchers routinely adopt them. Thus, our primary objectives in this study were to describe the psychometric properties of the RC scales, evaluate the extent to which they address problems associated with the standard clinical scales, and provide additional validity and interpretive data.

The empirical criterion-keying approach used by Hathaway and McKinley (1943) to construct the original clinical scales resulted in significant item overlap among the scales. In the MMPI–2, the average number of overlapping items per pair of the 10 standard clinical scales is 6.4 items (Greene, 2000; Helmes & Reddon, 1993), with notably higher overlap in certain scale pairs: 1–3 (20 items overlapping), 7–8 (17), 2–7 (13), 2–3 (13), 6–8 (13), 4–0 (11), 8–9 (11), 1–2 (10), 2–8 (10), 3–4 (10), and 4–8 (10). Redundancy of this magnitude is problematic for a number of reasons. First, item overlap results in artifactually high intercorrelations among scales and decreases their distinctiveness and discriminant validity (Helmes & Reddon, 1993; Tellegen et al., 2003). Item overlap also poses significant challenges to understanding the factor structure of MMPI/MMPI–2 scales (e.g., Horn, Wanberg, & Appel, 1973; Reddon, Marceau, & Jackson, 1982; Waller, 1999) and renders the meaning of individual scale scores unclear, especially when endorsement of common items elevates many scales simultaneously.

Another problematic by-product of empirical criterion-keying is that it often results in heterogeneous item content within scales (e.g., Greene, 2000; Helmes & Reddon, 1993; Tellegen et al.,

Leonard J. Simms, Department of Psychology, University at Buffalo, The State University of New York; Alex Casillas, Lee Anna Clark, and David Watson, Department of Psychology, University of Iowa; Bradley N. Doebbeling, Roudebush Veterans' Affairs Medical Center, Indianapolis, IN; Regenstrief Institute for Health Care; and Indiana University School of Medicine.

A previous version of this article was presented at the annual meeting of the Society for Personality Assessment, Miami, FL, March 2004. This work was supported by grants from the Department of Defense (DAMD17-97-1-7355) and the University of Minnesota Press. We appreciate the contributions of the Iowa Gulf War Study Group, the Carl E. Seashore Psychology Training Clinic, research assistants who helped with data collection and entry, and participants who provided the data included in this study.

Correspondence concerning this article should be addressed to Leonard J. Simms, Department of Psychology, Park Hall 218, University at Buffalo, State University of New York, Buffalo, NY 14260. E-mail: ljsimms@buffalo.edu

2003). The resulting multidimensionality of the clinical scales poses important challenges to accurate interpretation, particularly for scale scores of moderate elevation (i.e., *T* scores between 60 and 70). For example, how is one to interpret a *T* score of 65 on Scale 3? Greene (2000, p. 139) reported that "the 60 items of Scale 3 can be classified into two general categories: items reflecting specific somatic symptoms . . . and items that show that the client considers himself or herself well socialized and well adjusted." Presumably, endorsing many items from *both* of these facets is most consistent with the responses of the original criterion group. However, a *T* score of 65 can be reached by endorsing items from only one of the two facets (i.e., *either* somatic complaints *or* perceptions of good psychological adjustment) or by endorsing a mixture of items from both facets. Thus, proper interpretation of Scale 3 depends on the particular items endorsed. Similar heterogeneity exists for almost all of the clinical scales.[1]

Development of the RC Scales

Tellegen et al. (2003) reported that they adopted a flexible test construction strategy that permitted them to identify and measure the meaningful and distinct components of each standard clinical scale. They described four stages of test construction: (a) creation of a Demoralization scale that "measures a broad, emotionally colored variable that underlies much of the variance common to the MMPI–2 Clinical Scales" (Tellegen et al., 2003, p. 11), (b) identification of distinct "core" components for each clinical scale, (c) development of "seed scales" to tap each core component using items from the standard clinical scales, and (d) derivation of a final set of RC scales composed of items from the entire MMPI–2 item pool.

Tellegen et al. (2003) based the Demoralization scale on previous studies of the hierarchical structure of affect (Tellegen, Watson, & Clark, 1999; Watson & Tellegen, 1985), which suggest that affect can be assessed at different levels of specificity. In particular, they argued that the nonspecific component of the MMPI–2 clinical scales could be modeled as the highest level of the affect hierarchy, a *pleasantness—unpleasantness* dimension, which reflects general hedonic valence. By extracting this variance as an independent scale, they hoped to reduce the artificial covariation among—and therefore highlight the distinctiveness of—the remaining clinically meaningful dimensions of the clinical scales. They constructed a "seed" scale through a series of item-level principal-components analyses of Scales 2 and 7, identifying items that either (a) loaded on the first unrotated component or (b) loaded oppositely on both orthogonally extracted components of Positive Emotionality and Negative Emotionality. They constructed the final 24-item Demoralization scale (RCd) by identifying items from the rest of the MMPI–2 item pool that maximized the convergent and discriminant validity of the scale.

Tellegen and colleagues next identified the "core" component of each clinical scale and created a seed scale for each. For each clinical scale, they conducted principal-components analyses of the scale's items plus the Demoralization items discussed above and extracted as many orthogonal components as were necessary to yield a Demoralization component (which emerged as the first principal component in all analyses) and a component with content that appeared psychologically coherent and distinctive from all other core components. Of the eight clinical scales for which they developed RC scales, the core component appeared as the second component for five scales—Scales 1, 2, 7, 8, and 9—and the third component for Scales 3, 4, and 6.[2,3] They then constructed seed scales for each core component through an iterative process in which they identified face valid items that maximized the internal consistency of and distinctiveness among the seed scales (i.e., no item overlap).

Finally, Tellegen et al. (2003) completed the RC scales by identifying additional items from the entire MMPI–2 item pool that maximized the convergent and discriminant validity of each scale. In addition to RCd, these analyses yielded eight RC scales whose names reflect the core components assessed by each: Somatic Complaints (RC1), Low Positive Emotions (RC2), Cynicism (RC3), Antisocial Behavior (RC4), Ideas of Persecution (RC6), Dysfunctional Negative Emotions (RC7), Aberrant Experiences (RC8), and Hypomanic Activation (RC9). Notably, seven of these scales represent face valid derivatives of the content of the original clinical scales, but the restructuring markedly changed the content of Scale 3 by including only items tapping cynical views of the world, because the somatic items that originally composed a significant portion of Scale 3 were redundant with the core component of RC1. Therefore, it is quite likely that RC3 and the original Hysteria scale will have different validity correlates.

Evidence of Psychometric Properties

Tellegen et al. (2003) compared the psychometric properties of the RC scales to those of the standard clinical scales. With regard to internal consistency, they found the RC scales to be equal or superior to the standard clinical scales. Median coefficient alphas ranged from .76 (for the normative sample of men and women) to .86 (male psychiatric patients) for the RC scales, whereas the comparable median alphas generally were lower for the standard clinical scales, ranging from .59 (for the normative sample of men) to .79 (female psychiatric patients). The RC scales also yielded higher temporal stability (median *r*s = .81 and .86 for men and women, respectively) in a portion of the normative sample over a short interval ($M = 9$ days) compared with the standard clinical scales (median *r*s = .74 and .68 for men and women, respectively). Tellegen et al. (2003) also reported data suggesting that, compared with the standard clinical scales, the RC scales (a) converged well with their parent scales (except for RC3), (b) were more distinctive, (c) yielded lower relations with RCd (Demoralization), and (d) generally demonstrated improved discriminant validity and

[1] Scale 7, which was developed using factor analytic techniques, is somewhat of an exception, although the scale is sufficiently long that its content is still quite broad.

[2] Tellegen et al. (2003) identified core components and seed scales for Scales 5 and 0 but ultimately did not produce RC scales. They argued that these scales do not represent basic dimensions of psychopathology.

[3] Tellegen et al. (2003) reported that the second principal components for Scales 3, 4, and 6 reflected content redundant with other core components. For example, the second component extracted from Scale 6 included items tapping Cynicism, which had already been identified as the core component of RC3. Thus, the third extracted component, Ideas of Persecution, served as the core component for the restructured version of Scale 6.

equivalent or superior convergent validity when correlated with variables conceptually related to the core components of the scales.

Thus, the initial reliability and validity data are promising, but replication and extension of these findings in independent samples are necessary. It is tempting to assume that the large body of empirical research on the standard clinical scales, especially their interpretive correlates, can be applied to these new conceptually matched scales. However, although several conference presentations and symposia have supported the psychometric properties of the RC scales (e.g., Arbisi, 2004; Osberg, Haseley, Kamas, & Henderson, 2004; Simms, Casillas, Clark, & Watson, 2004), no published studies of the RC scales have appeared in the literature to date.

However, the RC scales are not without their critics. Nichols (2004) (a) argued that the RC scales largely are redundant with other MMPI–2 scales (e.g., the content and PSY-5 scales), (b) criticized the methods and theoretical model of affect used to construct RCd, and (c) argued that the criteria against which Tellegen et al. (2003) compared and validated the standard and restructured clinical scales were insufficient to reflect the complexity and heterogeneity of the standard clinical scales, thereby putting them at a disadvantage in head-to-head comparisons with the RC scales. His first 2 points are notable and deserve further study, but they are beyond the scope of the present article. However, with regard to his last argument, Nichols wrote, "Just as the multidimensional clinical scales may possess advantages in the prediction of complex criteria such as psychiatric syndromes/diagnoses, unidimensional scales will quasi-always have an advantage over complex, syndromal scales in predicting to discrete correlates" (p. 10). He then added, "A far more informative demonstration of the comparative validities of the RC versus clinical scales would direct the RC scales to more appropriate but more challenging multivariate correlates such as SCID . . . diagnoses" (p. 10).

Clearly, more data are needed before the RC scales can be used and interpreted with confidence. Therefore, we investigated the psychometrics and construct validity of the RC scales in samples of psychology clinic clients and military veterans who completed the MMPI–2 and other measures. Our study extends previous work by reporting data from additional sample types and comparing the construct validity of the restructured and standard clinical scales in relation to measures of psychopathology and personality disorder (PD) not previously reported. Moreover, we attempt to address Nichols' (2004) third concern by presenting comparative validity evidence against both unidimensional and multidimensional criteria.

Method

Participants and Procedures

Psychology clinic sample. Charts from adult patients seen between 1994 and 2002 at the Carl E. Seashore Psychology Training Clinic—an outpatient psychology clinic in the Department of Psychology at the University of Iowa—were reviewed and rated on a number of variables. Chart-documented reasons for referral primarily included (a) psychotherapy for anxiety, depression, distress, or family/relationship problems (43.5%); (b) court, custody, and disability evaluations (29.5%); and (c) assessments related to learning and attentional problems (12.3%). We selected charts from clients who previously had completed both the MMPI–2 and the Schedule for Nonadaptive and Adaptive Personality (SNAP; Clark, 1993). Of the 777 eligible individuals, 285 (36.7%) had completed these measures. These participants averaged 32.3 years of age (SD = 10.5 years) and were primarily female (56.8%), Caucasian (71.0%), single (45.3%), and employed (52.1%).[4] Current, chart-documented Axis I diagnoses included mood (23.9%), adjustment (7.1%), anxiety (6.8%), or substance-related (6.6%) disorders. Additionally, Axis II diagnoses included personality disorder not otherwise specified (9.2%), as well as PDs from Clusters B (7.4%), C (5.6%), and A (3.2%).

Military veteran sample. This sample is part of the Iowa Gulf War Study (Iowa Persian Gulf Study Group, 1997), which began as a telephone interview that assessed a broad range of health-related characteristics among veterans who served in the military during the first Gulf War. Participants were eligible for the study if (a) they served active duty or were activated National Guard or U.S. Army Reserve during the first Gulf War between August 2, 1990, and July 31, 1991, and (b) enlistment records indicated the state of Iowa as participants' state of residence. The original sampling base consisted of 29,010 eligible personnel. From this pool, a total of 4,886 individuals (16.8%) were selected randomly from four deployment and duty status domains (i.e., Deployed Active Duty, Deployed National Guard/Reserve, Nondeployed Active Duty, and Nondeployed National Guard/Reserve) and further stratified on the basis of military branch, rank, age, sex, and ethnicity. Samples were drawn proportionally from each of the 64 strata, with oversampling of small strata. Of the 4,886 individuals sampled, 3,695 (75.6%) participated from September 1995 to May 1996; these participants are henceforth referred to as *Phase I participants* (see Doebbeling et al., 2002; Iowa Persian Gulf Study Group, 1997; Simms, Watson, & Doebbeling, 2002, for more details regarding Phase I study methodology).

Subsequently, Phase I participants who met a priori criteria for any of three symptom constellations—cognitive dysfunction, chronic widespread pain, or major depression—were oversampled to participate in a second phase of data collection; Phase II also included control participants who were asymptomatic at Phase I. Of the 3,695 Phase I participants, 1,059 (28.7%) met these criteria and were invited to participate in a full-day assessment. Of these, 602 (56.8%) participated in Phase II data collection between March 1999 and July 2002. For the present study, we limited analyses to participants who completed the MMPI–2, SNAP, and the Structured Clinical Interview for Axis I *DSM–IV* Disorders (SCID; First, Spitzer, Gibbon, & Williams, 1997). The resulting 567 participants averaged 39.2 years of age (SD = 9.0 years) and were mostly male (87.7%), Caucasian (97.2%), married or cohabiting (73.6%), and deployed during the Gulf War (73.0%). In addition to the MMPI–2, SNAP, and SCID, the Phase II assessment included additional surveys and records that were not the focus of the present study.

Measures

MMPI–2. The MMPI–2 (Butcher et al., 2001) consists of 567 statements to which participants respond using a true–false format. In addition to the 10 standard clinical scales, the MMPI–2 includes a variety of validity, content, and supplementary scales that commonly are scored. Thousands of studies have appeared supporting the validity and clinical utility of the MMPI and MMPI–2 (e.g., Butcher & Williams, 2000; Graham, 2000; Greene, 2000); however, as discussed above, independent research on the RC scales has yet to be published.

SNAP. The SNAP (Clark, 1993) is a factor analytically derived, self-report instrument designed to assess trait dimensions important in the domain of PDs. The SNAP contains 375 items and uses a true–false format.

[4] Ethnicity data for 56 (19.8%) psychology clinic participants were not available from their charts. However, on the basis of census data for the catchment area of the clinic, most of these individuals are likely to be Caucasian.

The core of the SNAP is 12 trait scales that assess specific or primary traits (Mistrust, Manipulativeness, Aggression, Self-Harm, Eccentric Perceptions, Dependency, Exhibitionism, Entitlement, Detachment, Impulsivity, Propriety, and Workaholism) and three temperament scales that reflect individual differences in the broad dimensions of Negative Temperament, Positive Temperament, and Disinhibition. These scales are internally consistent (median coefficient $\alpha = .83$, range = .72–.92) and have acceptable retest reliabilities (median $r = .87$, range = .75–.90, mean interval = 49 days; Clark, 1993; Clark, Simms, Wu, & Casillas, in press). In addition, the SNAP's 15 trait and temperament scales have shown good convergent and discriminant relations with other relevant self-report measures, including measures of the three- and five-factor models of personality, state and trait mood, symptoms of depression and anxiety, and the MMPI–2 (see Clark, 1993, and Clark et al., in press, for details).

The SNAP also includes diagnostic scales keyed to the PDs in the revised third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM–III–R*; American Psychiatric Association, 1987). The PD scales can be scored categorically or dimensionally, with the latter method yielding superior reliability and validity coefficients. Scored dimensionally, the PD scales are reliable, with median alphas ranging from .72 to .82 in a variety of patient and nonpatient samples. In addition, the PD scales have shown moderate to good convergence with interview ratings. In a mixed patient sample ($N = 89$), for example, Clark (1993) reported a median correlation of .55 (range = .38–.85) between the SNAP PD scales and the corresponding scales of the Structured Interview for *DSM–III–R* Personality (SIDP—R; Pfohl, Blum, Zimmerman, & Stangl, 1989).

SCID. The SCID (First et al., 1997) is a commonly used structured interview for assessing current and lifetime psychiatric disorders on the basis of *DSM–IV* diagnostic criteria. Previous studies have suggested that the SCID can be administered with adequate to excellent interrater (Zanarini et al., 2000) and test–retest (Williams et al., 1992) reliabilities. In the present study, the SCID was administered only to the military veteran sample. Interviewers received significant training in use and administration of the SCID. The interrater reliability of SCID diagnoses was compared on three separate occasions using audiotapes from 30 randomly selected cases, with kappa coefficients consistently in excess of .80.

Data Analysis and Results

We compared the RC scales to the standard clinical scales with regard to descriptive statistics, internal consistency, covariance structure, and convergent and discriminant validity. For most analyses, we present cross-validation evidence from both samples. Because of the large number of comparisons and statistical tests associated with each primary analysis, we set the p value for significance conservatively to .01 for all analyses. The presentation of descriptive statistics is notably limited in the RC scales manual; thus, we present T score descriptive statistics for the restructured and standard clinical scales (without K correction) in Table 1 for both of our samples not only to aid interpretation of our primary results but also to add to the limited body of descriptive data available for the RC scales.

Internal Consistency

Reliability analyses were conducted to compare internal consistencies across the two scale sets. Because scales constructed with factor analytic methods generally are more homogeneous and efficient than those constructed using empirical criterion keying, we expected the RC scales to yield internal consistency coefficients at least as large as those of the standard clinical scales despite using, on average, 60% fewer items. Indeed, alpha coefficients and average interitem correlations (AICs), presented in

Table 1
Descriptive Statistics for the Standard and Restructured Clinical (RC) Scales

	Psychology clinic ($N = 285$)		Military veterans ($N = 567$)	
Scale	*M*	*SD*	*M*	*SD*
L	52.8	11.6	52.1	9.7
F	59.8	17.9	54.4	14.3
K	49.1	10.6	48.8	10.7
1. Hypochondriasis	58.0	14.6	59.8	14.1
2. Depression	62.8	15.5	56.5	14.0
3. Hysteria	59.3	15.3	56.2	13.5
4. Psychopathic Deviate	62.6	13.6	54.3	12.3
6. Paranoia	61.6	14.8	53.7	12.6
7. Psychasthenia	59.8	15.3	54.9	14.1
8. Schizophrenia	60.1	16.5	55.1	14.6
9. Hypomania	52.4	11.2	51.9	10.7
10. Social Introversion	56.3	11.6	53.9	12.0
RCd: Demoralization	60.0	14.4	53.7	13.1
RC1: Somatic Complaints	57.6	14.6	58.7	14.3
RC2: Low Positive Emotions	58.1	15.2	51.4	11.8
RC3: Cynicism	51.8	9.4	53.9	10.1
RC4: Antisocial Behavior	56.0	11.3	51.1	10.0
RC6: Ideas of Persecution	56.9	12.6	52.8	11.1
RC7: Dysfunctional Negative Emotions	54.4	13.4	51.5	12.4
RC8: Aberrant Experiences	53.4	12.7	52.1	11.1
RC9: Hypomanic Activation	49.4	11.3	50.8	10.1

Note. Scores are T score conversions (i.e., $M = 50$, $SD = 10$), without K corrections. RCd = Demoralization scale.

Table 2, indicated that the eight substantive RC scales (i.e., those other than RCd) were equivalently to slightly more internally consistent (median alphas = .83 and .77 in the psychology clinic and military veteran samples, respectively) than their standard clinical scale counterparts (median alphas = .78 and .76, respectively).

A closer examination of the AIC values revealed two scale groupings. For Scales 2, 3, 4, 6, and 9, AICs were close to zero (mean AICs = .05 [*SD* = .01] and .04 [*SD* = .01] in the patient and veteran samples, respectively), whereas the corresponding RC scale AICs were considerably larger (mean AICs = .18 [*SD* = .04] and .15 [*SD* = .03], respectively). In contrast, the AICs for Scales 1, 7, and 8 (*M*s = .20 [*SD* = .03] and .17 [*SD* = .04], respectively) were roughly equivalent to those for their RC scale counterparts (*M*s = .23 [*SD* = .03] and .19 [*SD* = .04], respectively). The restructured versions of Scales 6 and 9 yielded the largest reliability gains: Alphas increased from approximately .60 to .80 and AICs increased from approximately .03 to .15. Thus, as expected, the RC scales largely increased measurement efficiency, improving upon the reliability of the standard clinical scales while using substantially fewer items. Of course, such results likely reflect differences not only in the methods of scale construction but also in the nature of the constructs being measured by these two sets of scales. The factor analytic roots of the RC scales are well suited for producing unidimensional measures of relatively discrete constructs, whereas the empirical methods used to develop the standard clinical scales resulted in heterogeneous measures of complex constructs.

Internal Structure

Intercorrelations between and within these groups of scales are presented in Table 3, separately by sample. Given the goals and methods that guided Tellegen et al.'s (2003) construction of the RC scales, we expected these correlations to reveal (a) generally good correspondence between the RC scales and their standard scale counterparts and (b) better discriminant validity among the RC scales. First, the convergence between matching scales generally was good. In the psychology clinic sample, seven of the eight substantive RC scales were correlated most strongly with their corresponding standard clinical scales. In the military veterans, the same held true for six of eight RC scales. In contrast, RC3 and Scale 3, which correlated at −.12 and −.14 across samples, clearly did not converge. It is important to recall that all of these relations represent part–whole correlations to varying degrees because of item overlap (amount of overlap is listed in Table 2). Nevertheless, with the exception of RC3, there was reasonable convergence among matched pairs of scales.

In addition, intercorrelations among the restructured and standard clinical scales indicated that the RC scales were somewhat more distinct from one another (mean *r* = .42 in both samples) than were the standard clinical scales (mean *r*s = .60 and .55 in the patient and veteran samples, respectively). Of the 28 corresponding discriminant correlations in each sample, 16 (57%) and 14 (50%), respectively, were significantly lower among the RC scales than the standard clinical scales, *z*s ranged from 2.0 to 13.6; another 10 (36%) and 9 (32%) correlations, respectively, were not significantly different between scale sets; and only a small number, 2 (7%) and 5 (18%), respectively, were significantly higher among the RC scales; *z*s ranged from 2.4 to 7.9. Again, these findings are not surprising given the different methods used to construct these sets of scales. Notably, however, some specific scale-pairings that yielded very high correlations in the original scales showed better discriminant validity in the restructured set. Scales 7 and 8, for example, correlated .91 in both samples, whereas the correlation between RC7 and RC8 was significantly lower (*r*s = .65 and .62 in the patient and veteran samples, respectively, *z*s = 9.0 and 13.6). Further, although Scales 2 and 7 correlated at .79 and .76 across samples, the restructured versions were significantly less correlated (*r*s = .55 and .47 in the patient and veteran samples, respectively, *z*s = 5.5 and 8.5). Finally, the .72 correlation between

Table 2
Scale Lengths and Internal Consistency Reliability Coefficients for the Standard and Restructured Clinical (RC) Scales

				Cronbach's alpha				Average interitem correlations			
	Number of items			Standard scales		RC scales		Standard scales		RC scales	
Scales	Standard	RC	Overlap (%)[a]	PC	MV	PC	MV	PC	MV	PC	MV
RCd[b]		24				.94	.93			.39	.36
1/RC1	32	27	20 (74)	.90	.89	.89	.87	.22	.20	.23	.20
2/RC2	57	17	8 (47)	.81	.79	.83	.77	.07	.06	.22	.16
3/RC3	60	15	5 (33)	.73	.69	.76	.77	.04	.04	.17	.18
4/RC4	50	22	9 (41)	.72	.70	.77	.77	.05	.04	.13	.13
6/RC6	40	17	13 (77)	.61	.53	.81	.73	.04	.03	.20	.14
7/RC7	48	24	8 (33)	.93	.92	.89	.87	.22	.19	.25	.22
8/RC8	78	18	10 (56)	.94	.92	.82	.75	.17	.13	.20	.14
9/RC9	46	28	8 (29)	.62	.61	.84	.79	.03	.03	.16	.12
M	51	21	10 (49)	.78	.76	.83	.79	.10	.09	.20	.16

Note. *N*s = 285 and 567 for psychology clinic (PC) and military veteran (MV) samples, respectively. Standard = original clinical scales; RCd = Demoralization scale.
[a] Percentage of RC scale items scored on the corresponding standard clinical scale. [b] RCd has no standard clinical scale equivalent and is not included in column mean calculations.

Table 3
Intercorrelations Among the Standard and Restructured Clinical (RC) Scales

Scale	RCd	RC1	RC2	RC3	RC4	RC6	RC7	RC8	RC9	1	2	3	4	6	7	8	9
RCd: Demoralization	—	**.61**	**.67**	**.51**	.40	**.51**	**.79**	**.55**	.34	**.64**	**.75**	.32	**.71**	**.55**	**.90**	**.87**	.34
RC1: Somatic Complaints	**.58**	—	**.53**	.31	.29	.34	**.57**	.46	.27	***.95***[a,b]	**.69**	**.68**[a]	.48	.42	**.69**	**.71**	.31
RC2: Low Positive Emotions	**.72**	.49	—	.23	.26	.31	.47	.26	−.03	**.57**	***.77***[a,b]	.37	**.54**	.46	**.62**	**.62**	−.05
RC3: Cynicism	.39	.29	.19	—	.38	.49	**.56**	**.50**	**.54**	.33	.25	*−.14*	.35	.03	**.54**	**.57**[b]	.43
RC4: Antisocial Behavior	.38	.32	.26	.22	—	.33	.40	.33	.48	.29	.18	.05	***.61***[a,b]	.32	.41	.49	.41
RC6: Ideas of Persecution	.43	.49	.27	**.50**	.32	—	**.57**	**.63**	.39	.35	.32	.08	**.51**	***.57***[a]	**.53**	**.64**[b]	.45
RC7: Dys. Negative Emotions	**.78**	**.61**	**.55**	.46	.39	**.56**	—	**.62**	.49	**.57**	**.55**	.13	**.55**	**.46**	***.87***[a,b]	**.81**	.42
RC8: Aberrant Experiences	**.50**	**.59**	.26	.41	.35	**.65**	**.65**	—	**.51**	.46	.36	.18	.49	.45	**.63**	***.74***[a,b]	**.59**
RC9: Hypomanic Activation	.36	.32	.02	.45	.47	.45	**.51**	**.54**	—	.27	−.01	−.07	.28	.16	.44	.47	***.73***[a,b]
1. Hypochondriasis	**.64**	***.95***[a,b]	**.58**	.26	.32	.44	**.60**	**.54**	.26	—	**.73**	**.70**	**.51**	.41	**.71**	**.72**	.30
2. Depression	**.79**	**.64**	***.80***[a,b]	.21	.19	.33	**.61**	.33	.03	**.73**	—	**.56**	**.56**	.48	**.76**	**.70**	.05
3. Hysteria	.46	**.72**[a]	.48	*−.12*	.16	.18	.30	.30	.01	**.78**	**.62**	—	.38	.37	.34	.37	.10
4. Psychopathic Deviate	**.73**	**.52**	**.56**	.36	***.65***[a,b]	**.51**	**.58**	.44	.40	**.55**	**.61**	.45	—	**.63**	**.66**	**.72**	.42
6. Paranoia	**.61**	**.57**	.49	.16	.35	***.70***[a,b]	**.59**	**.54**	.34	**.58**	**.58**	.48	**.66**	—	**.53**	**.59**	.27
7. Psychasthenia	**.92**	**.68**	**.70**	.44	.38	**.53**	***.87***[a,b]	**.62**	.48	**.72**	**.79**	.48	**.70**	**.67**	—	**.91**	.41
8. Schizophrenia	**.84**	**.75**	**.64**	.48[b]	.47	**.66**	**.84**[a]	***.76***[b]	**.54**	**.74**	**.70**	.48	**.72**	**.71**	**.91**	—	**.52**
9. Hypomania	.35	.37	−.03	.46	.43	.45	.44	**.59**	***.79***[a,b]	.30	.04	.09	.44	.29	.43	**.53**	—

Note. Intercorrelations for the psychological clinic ($N = 285$; rs $> .16$ are significant, $p < .01$) and military veteran ($N = 567$; rs $> .11$ are significant, $p < .01$) samples are presented below and above the diagonal, respectively. Correlations $\geq .50$ are presented in boldface. Convergent correlations are italicized. RCd = RC Demoralization scale; Dys. = dysfunctional.
[a] Highest cross-correlation of each original scale, separately by sample. [b] Highest cross-correlation of each RC scale, separately by sample.

Scales 4 and 8 in both samples was substantially reduced between RC4 and RC8 (rs = .35 and .33 in the patient and veteran samples, respectively, zs = 6.8 and 10.0). It is interesting to note that two restructured scales—RC3 and RC9—did not yield the same pattern of greater discriminant validity. In fact, of the discriminant correlations that were not significantly lower among the RC scales, 10 of 12 (83%) and 11 of 14 (79%) in the patient and veteran samples, respectively, involved RC3 and/or RC9. Thus, restructuring significantly improved discriminant validity among many, but not all, scales.

Although the RC scales on the whole were less intercorrelated than the standard clinical scales, they still included an appreciable nonspecific component, as evidenced by their correlations with RCd: Mean rs between RCd and the other RC scales were .56 and .54 in the patient and veteran samples, respectively. Most notably, RC7 correlated at .78 and .79, respectively, with RCd, suggesting that these two scales tap highly overlapping constructs. Moreover, RC7 and RC2 correlated at .55 and .47, respectively, which is somewhat higher than we would expect given the theoretical considerations that guided their development (i.e., the assumption of relative independence between positive affect and negative affect).

Personality Correlates

The standard clinical scales have accrued over 50 years of evidence supporting their validity and providing an actuarial foundation from which to interpret scores and profiles. Given their newness, the RC scales have no such research foundation; thus, we present the following data to compare and contrast the construct validity of these two sets of scales. Table 4 includes correlations in both samples of the two sets of MMPI–2 scales with the trait and PD scales of the SNAP. To aid the reader in interpreting the large number of correlations in this table, we have marked the correlations in accordance with Cohen's (1988) recommendations for evaluating the strength of correlational effects: Medium effects (rs between .30 and .49) are italicized, and large effects (rs $\geq .50$) are designated with boldface type.

Several general features of these data are notable. First, the overall magnitude of the correlations in this table suggests that both the restructured and standard clinical scales continue to relate nonspecifically to scales reflecting negative affectivity or general neuroticism (e.g., the first 7 scales listed for the SNAP, which usually form such a factor), but to different degrees. Of the 112 correlations presented in this part of the table for the standard clinical scales, 44 (39%) and 40 (36%) would be considered medium and large effects, respectively, using Cohen's criteria. In contrast, of the 112 correlations presented in this part of the table for the RC scales, 59 (53%) and 30 (27%) would be considered medium and large effects, respectively. Thus, although both sets of scales continue to be somewhat nonspecific, the RC scales yielded fewer large effects and more medium effects relative to the standard clinical scales, suggesting that the restructuring was somewhat successful in reducing to a more moderate level the nonspecific variance embodied in these scales.

A second notable point from these data are that, for most of the RC scales, the pattern of convergent and discriminant correlations is somewhat comparable to that computed with the standard clinical scales. As an index of validity comparability within each scale pairing (i.e., Scale 1 with RC1, Scale 2 with RC2, etc.), we correlated the columns of correlations. Results revealed that all but one of the RC scales yielded moderately to highly stable validity patterns (i.e., rs $> .80$); notably, RC3 yielded a substantially different validity pattern compared with Scale 3 ($r = .36$, 95% confidence interval = [.32, .40]), suggesting that removal of the somatic items from Scale 3 left a core component of Cynicism that has markedly different correlates from the original Scale 3. For example, RC3's highest SNAP correlate was Mistrust (a scale tapping interpersonal suspiciousness and general cynicism; rs =

Table 4
Schedule for Nonadaptive and Adaptive Personality (SNAP) Correlates of the Standard and Restructured Clinical (RC) Scales

		MMP1–2 Scale																
SNAP Scale	Sample	RCd	1	RC1	2	RC2	3	RC3	4	RC4	6	RC6	7	RC7	8	RC8	9	RC9
Trait and temperament scales																		
Negative Temperament	PC	**.79**	**.59**	**.56**	**.65**	**.55**	*.39*	*.36*	**.61**	*.36*	**.57**	*.41*	**.82**	**.81**	**.74**	*.48*	*.36*	*.48*
	MV	**.75**	**.57**	**.56**	**.57**	*.48*	.25	*.42*	**.54**	*.35*	*.47*	*.45*	**.79**	**.79**	**.72**	*.48*	*.34*	*.42*
Mistrust	PC	**.51**	*.43*	*.44*	*.38*	*.32*	.11	**.64**	**.57**	*.39*	**.50**	**.64**	**.57**	**.60**	**.64**	*.49*	*.45*	*.47*
	MV	**.68**	*.45*	*.43*	*.43*	*.42*	.08	**.67**	**.58**	*.42*	*.42*	**.67**	**.67**	**.66**	**.71**	**.56**	*.43*	*.45*
Manipulativeness	PC	**.51**	.29	.29	.25	.24	.14	.29	**.50**	*.49*	*.34*	*.35*	**.53**	**.53**	**.56**	*.42*	**.58**	**.62**
	MV	*.36*	.25	.22	.11	.19	−.05	*.46*	*.32*	*.45*	.12	*.36*	*.39*	*.42*	*.44*	*.38*	*.43*	**.57**
Aggression	PC	*.43*	*.33*	*.36*	.27	.27	.12	**.33**	**.50**	*.49*	*.38*	*.37*	*.46*	**.51**	**.51**	*.32*	*.42*	**.56**
	MV	*.42*	*.32*	*.31*	.22	*.31*	.05	*.42*	*.42*	*.48*	.29	*.39*	*.46*	**.50**	*.49*	*.34*	*.37*	**.55**
Self-Harm	PC	**.74**	**.53**	**.50**	**.64**	**.67**	*.38*	*.35*	**.66**	*.41*	**.54**	*.41*	**.72**	**.61**	**.72**	*.45*	.29	*.31*
	MV	**.69**	*.45*	*.44*	*.49*	**.57**	.26	*.32*	**.62**	*.39*	**.52**	*.43*	**.58**	**.53**	**.62**	*.35*	.25	.23
Eccentric Perceptions	PC	*.49*	*.42*	*.47*	*.32*	.27	.23	*.37*	*.45*	*.33*	*.45*	**.51**	**.58**	**.57**	**.67**	**.74**	**.56**	**.54**
	MV	**.50**	*.39*	*.38*	*.31*	.20	.19	*.41*	*.44*	*.30*	*.38*	**.54**	**.54**	**.51**	**.61**	**.67**	**.52**	*.45*
Dependency	PC	**.51**	*.30*	.29	*.41*	*.36*	.21	.18	*.35*	.15	*.32*	.23	**.50**	*.49*	*.43*	.26	.20	.22
	MV	*.38*	.24	.24	.28	.25	.10	.15	.23	.09	.21	.22	*.40*	*.38*	*.32*	.21	.09	.13
Positive Temperament	PC	−.51	*−.37*	−.27	**−.62**	**−.64**	*−.33*	−.04	*−.33*	−.10	−.21	−.03	*−.42*	−.27	*−.32*	.00	.22	.22
	MV	*−.46*	*−.37*	*−.32*	**−.56**	**−.68**	−.23	−.10	*−.34*	−.12	−.23	−.12	*−.38*	−.26	*−.37*	−.07	.20	.20
Exhibitionism	PC	.01	−.03	−.03	−.13	−.21	.02	.04	.14	.15	.07	−.01	.03	−.01	.05	.10	*.35*	*.44*
	MV	−.17	−.16	−.16	*−.32*	*−.37*	−.04	.00	.03	.11	.02	.00	−.13	−.13	−.07	.10	*.33*	*.38*
Entitlement	PC	−.12	−.07	−.05	−.19	*−.30*	−.14	.20	−.01	−.02	.01	.15	−.02	.02	.04	.15	*.35*	*.36*
	MV	−.08	−.05	−.05	−.19	−.29	−.11	.21	−.05	.03	−.03	.10	−.04	.01	.02	.20	.29	*.35*
Detachment	PC	*.48*	*.32*	.29	*.49*	**.55**	.10	.29	*.34*	.25	.28	.26	*.47*	*.44*	*.49*	.26	.02	.03
	MV	**.56**	*.35*	*.32*	**.51**	**.62**	.04	*.34*	*.32*	.24	.26	*.31*	**.50**	*.46*	**.52**	.25	−.01	.05
Disinhibition	PC	*.45*	.22	.21	.18	.22	.15	.20	*.48*	**.54**	*.30*	.25	*.45*	*.38*	*.46*	*.35*	**.52**	**.57**
	MV	*.35*	.22	.19	.14	.25	.02	*.36*	*.36*	*.45*	.18	.26	*.35*	*.31*	*.41*	*.30*	*.36*	*.48*
Impulsivity	PC	*.41*	.23	.24	.19	.21	.23	.10	*.42*	*.46*	.28	.17	*.41*	*.31*	*.40*	*.33*	*.44*	*.47*
	MV	.28	.22	.21	.16	.24	.10	.19	*.33*	*.31*	.21	.16	.29	.22	*.31*	.21	.24	*.31*
Propriety	PC	.02	.06	.07	.04	−.12	−.10	.24	−.04	−.14	.07	.20	.06	.16	.07	.13	.02	.05
	MV	.10	.03	.02	.02	−.12	−.10	.20	−.03	−.07	.00	.12	.16	.22	.07	.08	.12	.14
Workaholism	PC	.18	.21	.25	.12	.05	.04	.26	.17	.11	.23	*.31*	.23	*.30*	.27	.29	.25	*.30*
	MV	.18	.19	.20	.09	.00	.07	.25	.11	.08	.16	.28	.22	.25	.23	.20	.24	.24
Personality disorder scales																		
Paranoid PD	PC	**.54**	*.43*	*.45*	*.41*	*.36*	.09	**.65**	**.57**	*.39*	**.50**	**.63**	**.59**	**.62**	**.64**	*.49*	*.41*	*.46*
	MV	**.68**	*.43*	*.42*	*.44*	*.46*	.03	**.67**	**.56**	*.45*	*.40*	**.64**	**.67**	**.68**	**.71**	**.53**	*.39*	*.47*
Schizoid PD	PC	*.38*	.29	.27	*.40*	*.46*	.08	.29	.27	.21	.26	.29	*.37*	*.34*	*.41*	.23	.05	.00
	MV	**.50**	*.33*	*.31*	*.47*	**.57**	.07	*.33*	.29	.21	.21	.28	*.43*	*.37*	*.49*	.25	.02	.02
Schizotypal PD	PC	**.61**	*.47*	*.49*	*.46*	*.44*	.13	**.58**	**.55**	*.38*	**.54**	**.64**	**.69**	**.69**	**.75**	**.65**	*.45*	*.47*
	MV	**.69**	*.46*	*.45*	**.50**	**.52**	.09	**.56**	**.51**	*.36*	*.40*	**.61**	**.68**	**.67**	**.74**	**.59**	*.36*	*.36*
Antisocial PD	PC	*.35*	.24	.24	.12	.15	.10	.26	**.51**	**.74**	.25	.28	*.38*	*.37*	*.44*	*.38*	**.51**	**.54**
	MV	*.41*	.28	.27	.17	.27	.03	*.42*	**.50**	**.72**	.28	*.35*	*.42*	*.41*	**.50**	*.36*	*.41*	**.52**
Borderline PD	PC	**.73**	**.55**	**.54**	**.53**	**.51**	*.33*	*.37*	**.71**	**.61**	**.58**	*.46*	**.75**	**.74**	**.78**	**.54**	**.53**	**.59**
	MV	**.73**	**.54**	**.53**	*.49*	*.47*	.21	**.50**	**.65**	**.56**	*.49*	**.55**	**.74**	**.72**	**.76**	**.55**	*.48*	**.55**
Histrionic PD	PC	*.31*	.16	.17	.06	−.03	.05	.23	*.33*	*.30*	.28	.21	*.35*	*.34*	*.35*	*.32*	**.53**	**.65**
	MV	.14	.08	.09	−.09	−.16	.00	.23	.20	.25	.16	.20	.21	.21	.23	*.32*	*.46*	**.56**
Narcissistic PD	PC	*.44*	.29	.29	.23	.15	.10	*.42*	*.48*	*.33*	*.40*	*.41*	**.52**	**.50**	**.55**	**.43**	**.59**	**.68**
	MV	*.42*	.27	.24	.14	.07	.00	**.53**	*.36*	*.35*	.26	*.47*	*.46*	*.47*	**.51**	**.51**	**.54**	**.65**
Avoidant PD	PC	**.63**	*.43*	*.40*	**.60**	**.63**	.15	*.41*	*.42*	.24	*.41*	*.39*	**.63**	**.63**	**.61**	*.35*	.06	.11
	MV	**.67**	*.46*	*.43*	**.61**	**.67**	.08	*.42*	*.39*	.26	*.33*	*.40*	**.66**	**.62**	**.62**	*.32*	.03	.11
Dependent PD	PC	**.61**	*.41*	*.40*	**.53**	*.42*	.27	.27	*.47*	.22	*.44*	*.30*	**.64**	**.62**	**.57**	*.34*	.26	*.33*
	MV	**.56**	*.34*	*.33*	*.40*	*.33*	.10	*.30*	*.36*	.17	*.34*	*.37*	**.59**	**.58**	*.49*	*.34*	.19	.24
Obsessive–Compulsive PD	PC	*.37*	.28	*.30*	*.31*	.23	.10	*.36*	*.30*	.15	*.30*	*.35*	*.41*	*.47*	*.44*	*.32*	.26	*.33*
	MV	*.41*	*.31*	*.31*	.29	.23	.08	*.36*	.27	.16	.24	*.35*	*.43*	*.44*	*.41*	.28	.25	*.30*
Validity pattern convergence																		
r with RCd column			.98	.97	.96	.95	.84	.73	.93	.66	.93	.81	.99	.97	.98	.77	.15	.17
r between standard and RC columns			.99		.98		.36		.84		.84		.99		.87		.97	

Note. *N*s = 285 and 567 for the PC (psychological clinic) and MV (military veteran) samples, respectively. Correlations ≥ .30 and < .50 are italicized, and those ≥ .50 are presented in boldface. All *r*s > .16 (PC sample) and .11 (MV sample) are significant, $p < .01$. RCd = RC Demoralization scale; SNAP = Schedule for Nonadaptive and Adaptive Personality; PD = personality disorder.

.64 and .67 in the psychology clinic and military veteran samples, respectively), whereas Scale 3's highest correlates were Negative Temperament (a scale measuring broad individual differences in the tendency to experience negative emotions such as sadness, anxiety, and anger) in the clinic sample (r = .39) and Self-Harm (a scale tapping self-destructive behavior in the context of self-loathing) in the veteran sample (r = .26). Clearly, Scale 3 and RC3 are not interchangeable.

We used similar methods to examine the comparative discriminant validity of the restructured and standard clinical scales with respect to the SNAP. To do this, we correlated each column of correlations in Table 4 (i.e., Scale 1, RC1, Scale 2, RC2, etc.) with that of RCd. In this type of analysis, discriminant validity is supported to the extent that the scales yield validity patterns distinct from that of the nonspecific RCd scale. Results of these analyses—which appear at the bottom of Table 4—revealed two themes. First, all of the validity patterns, except for those of Scale 9/RC9, moderately to strongly tracked RCd (rs ranged from .66 to .99). Second, although five of eight RC scales yielded validity patterns no more distinctive than their parent clinical scales, three RC scales—RC4, RC6, and RC8—produced validity patterns that were significantly more distinct, zs ranged from 3.9–23.1, suggesting better discriminant validity for these scales.

Despite the correlational similarity of the general validity patterns, a number of more specific differences are apparent from Table 4 that suggest greater distinctiveness and conceptual clarity for some specific RC scales. For example, given the typical interpretation of Scales 6 and 8, we expected them as well as their restructured counterparts to correlate with SNAP scales reflecting interpersonal suspiciousness (i.e., SNAP Mistrust) and unusual perceptions and beliefs (i.e., SNAP Eccentric Perceptions), respectively. However, although both Scale 6 and RC6 were moderately to strongly correlated with several SNAP scales, RC6 was more strongly and specifically related to SNAP Mistrust, a scale with a clear conceptual match to RC6's core component, persecutory ideation. Similarly, whereas the original Scale 8 yielded moderate to large correlations with as many as 11 of the SNAP's 15 trait scales—including Negative Temperament (rs = .74 and .72), Self-Harm (rs = .72 and .58), Mistrust (rs = .57 and .67), and Eccentric Perceptions (rs = .67 and .61)—RC8 was more specifically related to Eccentric Perceptions (rs = .74 and .67, with all other rs $<$ .57), a construct centrally related to the core component of Scale 8 emphasized by Tellegen et al. (2003).

Likewise, given the typical interpretation of Scales 4 and 9, we expected them and their restructured partners to correlate with SNAP scales reflecting disinhibitory processes, externalizing behavior, and psychopathy. However, whereas RC4 yielded medium to large and theoretically consistent correlations with SNAP Disinhibition (a scale tapping broad differences in over- vs. undercontrolled behavior; rs = .54 and .45), Aggression (a scale measuring anger and its behavioral expression; rs = .49 and .48), and Manipulativeness (a scale assessing interpersonal exploitation; rs = .49 and .45), Scale 4 correlated most strongly with Self-Harm (rs = .66 and .62) and Negative Temperament (rs = .61 and .54), which are nonspecific scales of self-loathing and general distress, respectively. Similarly, whereas Scale 9's highest correlates were different across samples (Manipulativeness in the psychology clinic sample and Eccentric Perception in the military veteran sample), RC9's strongest correlate—Manipulativeness—was the same in both samples, suggesting either that RC9 may assess content beyond its core concept of Hypomanic Activation or that the particular type of Hypomanic Activation assessed by RC9 includes elements of interpersonal exploitation.

Finally, given previous research on the dimensional nature of depression (e.g., Clark & Watson, 1991), we predicted that Scales 2 and RC2 would be most strongly related to SNAP scales reflecting general neuroticism or negative affectivity as well as low positive affectivity. It is interesting to note that Scale 2 and RC2 produced subtle but potentially important validity differences. Although both correlated substantially with the Negative Temperament, Positive Temperament (negatively; a scale tapping broad differences in the experience of positive emotions), and Detachment (a measure of interpersonal and emotional distance) scales of the SNAP, RC2 yielded a conceptually cleaner pattern. Whereas Scale 2 correlated most strongly with Negative Temperament (rs = .65 and .62), RC2 was most related to Positive Temperament (rs = −.64 and −.68) and also produced a stronger relationship with Detachment than did the original Scale 2, reflecting the anhedonia and withdrawal that theoretically is central to RC2.

Diagnostic Correlates

Like the RC scales, the SNAP trait and temperament scales were constructed factor analytically, which may have increased their convergence artifactually and thus favored them in these comparisons. Therefore, in this section, we correlated the two sets of MMPI–2 clinical scales with diagnostic scales that are more likely to reflect the complex, heterogeneous diagnostic entities for which the standard scales were designed. We first present relations between the MMPI–2 and the dimensional PD scales of the SNAP (in the second part of Table 4). Again, at the specific scale level, the RC scales appeared to provide a somewhat more distinctive pattern of relations with the SNAP PD scales. For example, 6 of 8 standard clinical scales had their strongest relations with the Borderline PD scale of the SNAP. In contrast, only 2 of 8 of the substantive RC scales yielded the same pattern (3 of 9 if RCd is included). In addition, the 3 RC scales that yielded only moderate validity comparability appeared to provide greater discriminant validity than their standard scale counterparts. Within the RC8/Scale 8 validity comparisons, for example, Scale 8 correlated strongly with 7 of 10 PD scales in at least one sample, whereas the same was true for RC8 with only 4 of 10 PD scales. In particular, RC8 was most strongly related to the conceptually similar Schizotypal PD scale of the SNAP (rs = .65 and .59 in the clinic and veteran samples, respectively), whereas Scale 8 related nonspecifically across the range of PDs. Similarly, RC6 correlated more strongly with Paranoid PD (rs = .63 and .64) and Schizotypal PD (rs = .64 and .61) than did Scale 6 (rs ranged .40 to .54), which also correlated strongly with Borderline PD (rs = .58 and .49).

Scale 4 and RC4 provide a final interesting comparison. Scale 4 correlated most strongly with Borderline PD (rs = .71 and .65) and also with Paranoid (rs = .57 and .56), Schizotypal (rs = .55 and .51), and Antisocial PDs (rs = .51 and .50). RC4, in contrast, correlated strongly only with Antisocial PD (rs = .74 and .72) and Borderline PD (rs = .61 and .56), providing a cleaner and more predictable pattern of convergent relations. Thus, when validity pattern differences emerged, they generally were in the direction of greater discriminant validity for the RC scales.

We also examined convergent and discriminant relations between the MMPI–2 and several primary categories of Axis I disorders. Data from the SCID were reduced to form four broad categories of psychological disorder, scored dichotomously (i.e., present or absent). Participants were categorized as having a *depressive disorder* if they met threshold SCID criteria for major depressive disorder, dysthymia, or depressive disorder not otherwise specified. Similarly, those identified by the SCID as meeting criteria for generalized anxiety disorder, panic disorder, social phobia, specific phobia, obsessive–compulsive disorder, or posttraumatic stress disorder were categorized as having an *anxiety disorder*. We formed the *substance use disorder* category by combining those who met SCID criteria for abuse or dependence of alcohol or any other substance. Finally, the class *somatoform disorder* reflected individuals who met criteria for somatization disorder, hypochondriasis, or undifferentiated somatoform disorder. Data were collected and coded both for current and lifetime diagnoses.

Correlations between these disorder categories and the scales of the MMPI–2 appear in Table 5 and reveal a mixed picture. Ratings of depressive disorders, for example, were moderately related to six of eight original clinical scales (all but Scales 6 and 9), whereas only four RC scales—RCd, RC1, RC2, and RC7—were related to current and lifetime diagnoses of depressive disorders, respectively. Notably, RC3, RC4, and RC8 failed to predict depressive disorders, suggesting better discriminant validity for these scales as compared with their corresponding standard scales. The strongest predictors of current and lifetime depressive disorders were Scale 2 (rs = .44 and .37, respectively) and RCd (rs = .48 and .40, respectively). It is interesting to note that, whereas RCd and Scale 2 predicted current and lifetime depression equivalently (zs = 1.54 and 1.11, respectively; ps > .05), RC2 predicted current (r = .37) and lifetime (r = .31) depression significantly less well than both scales, with zs ranging from 2.26 to 2.73 (ps < .05).

Ratings of anxiety disorders correlated most strongly with the standard and restructured versions of Scales 1, 2, and 7, as well as RCd and the original Scale 8. Here, the comparison between Scale 7 and RC7 is the most conceptually interesting, with the results revealing that current and lifetime ratings of anxiety disorders were predicted more strongly by original Scale 7 (both rs = .37) than RC7 (both rs = .32), z = 2.51 (p < .01). For substance use disorders, RC4 correlated most strongly (rs = .25 and .38 for current and lifetime diagnoses, respectively), which represents a significant improvement over Scale 4 (rs = .14 and .20 for current and lifetime diagnoses, respectively), zs = 3.06 and 5.23, respectively (p < .01). Finally, ratings of somatoform disorders were marked most clearly by both Scale 1 and RC1, rs ranged from .27 to .29 across scales and timeframes, and they were not significantly different, zs = 0.79 and 0.78 for current and lifetime diagnoses, respectively.

Incremental Validity

An important question to ask in evaluating the RC scales is the extent to which they add incrementally to the prediction of the validity indicators. To examine this issue, we conducted a series of hierarchical regression analyses. In each analysis, we entered one of the trait or diagnostic variables presented in Tables 4, 5, or 6 as the dependent measure and entered as a block either the RC scales or the eight corresponding standard clinical scales as the predictor variables at Steps 1 and 2, respectively. In order to assess the incremental validity of both sets of clinical scales, we did the analyses two ways: (a) with the standard clinical scales entered at Step 1 and the RC scales entered at Step 2, which provided an index of the RC scales' incremental validity, and (b) with RC scales entered at Step 1 and the standard clinical scales at Step 2, which estimated the incremental validity of the standard clinical scales. Multiple regression procedures were used for predicting the dimensional SNAP scales, whereas logistic regression was used for predicting the binary SCID diagnoses.

The results, which are summarized in Tables 6 and 7, indicated that both the standard and restructured clinical scales added incrementally to the prediction of the dependent measures. However, the magnitude of the incremental effects was larger for the RC scales than for the standard clinical scales. For the analyses in which the SNAP trait and temperament scales were entered as

Table 5
Structural Clinical Interview for DSM–IV Disorders (SCID) Correlates of the Standard and Restructured Clinical Scales in the Military Veteran Sample

	MMPI–2 Scale																
SCID Diagnosis (BR)	RCd	1	RC1	2	RC2	3	RC3	4	RC4	6	RC6	7	RC7	8	RC8	9	RC9
Depressive disorders																	
Current (13.7%)	**.48**	**.40**	**.37**	**.44**	**.37**	**.30**	.20	**.36**	.15	.29	.26	**.43**	**.33**	**.42**	.22	.10	.10
Lifetime (29.1%)	**.40**	**.33**	**.33**	**.37**	**.31**	**.30**	.07	**.33**	.17	.27	.17	**.35**	.28	**.36**	.20	.09	.07
Anxiety disorders																	
Current (19.0%)	**.36**	**.33**	**.33**	**.37**	**.33**	.22	.15	.27	.12	.27	.24	**.37**	**.32**	**.35**	.25	.11	.07
Lifetime (24.5%)	**.35**	**.33**	**.34**	**.36**	**.30**	.22	.12	.27	.12	.27	.20	**.37**	**.32**	**.34**	.25	.10	.07
Substance use disorders																	
Current (4.4%)	.11	.11	.10	.02	.09	−.01	.15	.14	.25	.10	.11	.11	.11	.11	.09	.09	.11
Lifetime (56.0%)	.10	.05	.05	.01	.01	.01	.09	.20	**.38**	.11	.11	.09	.11	.11	.08	.17	.19
Somatoform disorders																	
Current (2.0%)	.15	.28	.29	.20	.14	.23	.09	.10	−.01	.09	.08	.19	.14	.18	.11	.05	.05
Lifetime (2.1%)	.15	.27	.28	.20	.14	.21	.09	.09	−.02	.09	.07	.18	.15	.17	.11	.04	.04

Note. N = 564. All rs ≥ .30 are presented in boldface. All rs > .11 are significant (p < .01). BR = disorder class base rate in present sample; RCd = Demoralization scale; RC = restructured clinical scale.

Table 6
Incremental Validity Regression Analyses With Schedule for Nonadaptive and Adaptive Personality (SNAP) Scales as Dependent Variables

	Military veterans (N = 567)				Psychology clinic (N = 285)			
	IncrV of RCs		IncrV of SCs		IncrV of RCs		IncrV of SCs	
Dependent variable	SCs R^2	RCs ΔR^2	RCs R^2	SCs ΔR^2	SCs R^2	RCs ΔR^2	RCs R^2	SCs ΔR^2
SNAP Trait and Temperament scales								
Negative Temperament	.643*	.049*	.675*	.016*	.687*	.082*	.746*	.023*
Mistrust	.577*	*.149**	.683*	.008	.535*	*.102**	.605*	.032*
Manipulativeness	.333*	.098*	.406*	.024*	.472*	.072*	.512*	.032
Aggression	.315*	*.149**	.452*	.013	.354*	*.133**	.459*	.028
Self-Harm	.470*	*.102**	.529*	.043*	.596*	.042*	.624*	.015
Eccentric Perceptions	.443*	.077*	.505*	.014	.521*	.086*	.594*	.012
Dependency	.183*	.025	.180*	.028	.261*	.037	.289*	.009
Positive Temperament	.423*	*.112**	.512*	.023*	.480*	.076*	.526*	.030
Exhibitionism	.268*	*.130**	.336*	.061*	.192*	*.172**	.301*	.063*
Entitlement	.141*	*.125**	.243*	.022	.180*	*.133**	.259*	.054*
Detachment	.443*	.085*	.463*	.065*	.420*	.058*	.392*	.086*
Disinhibition	.242*	*.124**	.352*	.013	.420*	*.109**	.499*	.029
Impulsivity	.136*	.070*	.188*	.018	.321*	*.110**	.401*	.033
Propriety	.099*	.079*	.159*	.018	.075*	*.147**	.186*	.036
Workaholism	.098*	.060*	.147*	.011	.124*	.055	.158*	.021
M of Trait and Temperament scales	.321	.096	.389	.025	.376	.094	.437	.034
SNAP Dimensional PD scales								
Paranoid PD	.597*	.096*	.685*	.009	.565*	.092*	.617*	.039*
Schizoid PD	.380*	.084*	.385*	.078*	.286*	.094*	.301*	.079*
Schizotypal PD	.594*	.050*	.623*	.022*	.657*	.036*	.655*	.038*
Antisocial PD	.395*	*.194**	.578*	.011	.415*	*.199**	.603*	.010
Borderline PD	.645*	.056*	.696*	.005	.676*	.081*	.743*	.014
Histrionic PD	.275*	*.108**	.351*	.032*	.350*	*.151**	.473*	.027
Narcissistic PD	.417*	*.116**	.517*	.015	.468*	*.112**	.518*	.062*
Avoidant PD	.594*	.071*	.611*	.054*	.582*	.055*	.576*	.062*
Dependent PD	.375*	.031*	.377*	.029*	.423*	.048*	.444*	.026
Obsessive–Compulsive PD	.203*	.046*	.245*	.004	.219*	.070*	.258*	.030
M of PD scales	.448	.085	.507	.026	.464	.094	.519	.039

Note. Analyses are hierarchical multiple regressions. ΔR^2 values $> .100$ are italicized. IncrV = incremental validity; RCs = restructured clinical scales; SCs = standard clinical scales; PD = personality disorder.
* $p < .01$.

dependent measures (see Table 6), the RC scales predicted, on average, an additional 9.6% and 9.4% of the variance above that predicted by the standard clinical scales in the veteran and clinic samples, respectively. In contrast, when the standard clinical scales were entered as the second block, they accounted for an additional 2.5% and 3.4% of the variance, respectively, above that predicted by the RC scales. Similarly, with the SNAP PD scales as the dependent measures, the RC scales predicted, on average, an additional 8.5% and 9.4%, respectively, of the variance above that predicted by the standard clinical scales, compared with incremental effects of 2.6% and 3.9%, respectively, for the standard clinical scales. Moreover, across all SNAP variable regression analyses, significant incremental RC scale effects far outnumbered incremental standard scale effects. Statistically, 96% (24 of 25) and 92% (23 of 25) of incremental RC scale effects were significant ($p < .01$) in the veteran and clinic samples, respectively, compared with 44% (11 of 25) and 40% (10 of 25), respectively, of the incremental standard clinical scale effects.

For the logistic regression analyses involving SCID diagnoses as dependent measures (see Table 7), the results revealed very little incremental validity for either set of scales. The only significant incremental advantages were for prediction of current and lifetime ratings of substance use disorders, with the RC scales adding significantly to the predictive ability of the original clinical scales, ΔR^2s = 3.5% and 10.5%, $\Delta\chi^2$s(9) = 21.5 and 67.2, respectively. Notably, the incremental effect of the RC scales was carried exclusively by RC4, Wald's χ^2s(1) = 13.8 and 48.2, respectively (both odds ratios = 1.36).

To further quantify the diagnostic accuracy of the scale sets in these logistic regressions, we conducted a series of receiver operating characteristic (ROC) analyses (see McFall & Treat, 1999, for a discussion of the use of ROC analyses in clinical assessments). We first computed logistic regression scores for each participant, predicting each disorder class in three ways: (a) using only the standard clinical scales as predictors, (b) using only the RC scales as predictors, and (c) using all scales as predictors. We then

Table 7
Incremental Validity Regression Analyses in the Military Veteran Sample With Structured Clinical Interview for DSM–IV Disorders (SCID) Diagnostic Classes as Dependent Variables

	IncrV of RCs		IncrV of SCs		ROC Area Under Curve		
Dependent variable	SCs R^2	RCs ΔR^2	RCs R^2	SCs ΔR^2	SCs	RCs	SCs+RCs
Depressive disorders							
Current	.193*	.020	.202*	.011	.86	.85	.86
Lifetime	.167*	.025	.177*	.014	.75	.77	.77
Anxiety disorders							
Current	.142*	.010	.144*	.008	.77	.77	.77
Lifetime	.143*	.011	.146*	.008	.75	.76	.76
Substance use disorders							
Current	.048*	.035*	.061*	.023	.79	.82	.85
Lifetime	.063*	*.105**	.161*	.007	.65	.74	.74
Somatoform disorders							
Current	.069*	.007	.067*	.009	.95	.92	.96
Lifetime	.062*	.012	.066*	.008	.91	.92	.94

Note. $N = 567$. Analyses are hierarchical logistic regressions. ΔR^2 values $> .100$ are italicized. SCID = Structured Clinical Interview for *DSM–IV* Disorders; IncrV = incremental validity; RCs = restructured clinical scales; SCs = standard clinical scales; ROC = receiver operating characteristic.
* $p < .01$.

subjected the regression scores to nonparametric ROC analyses and calculated area under the curve (AUC) statistics—which provide an estimate of the overall accuracy of a given assessment procedure that is not confounded by changing cut scores or base rates—for each set of regression scores. The AUC values, which appear in Table 7, provided further evidence that the standard and restructured clinical scales were roughly equivalent in their ability to predict broad diagnostic class membership. Within each disorder class, the 95% confidence intervals around the AUC values were highly overlapping across scales sets. Mirroring the results presented above, the only exception was for prediction of lifetime substance use disorders, with AUC values of .74 and .65 for the restructured and standard clinical scales, respectively.

Discussion

The results of this study broadly suggest that the RC scales represent a somewhat successful attempt to improve on weaknesses often associated with the original clinical scales while clarifying the core constructs of each. A significant strength of the RC scales is their efficiency advantage over their clinical scale counterparts. However, greater efficiency may not be advantageous to the extent that the multidimensional character of the standard clinical scales has predictive utility and has been sacrificed in the process (Nichols, 2004). Thus, an important question we examined was the extent to which the standard and restructured clinical scales provided comparable validity evidence across criteria sets varying in complexity, ranging from the unidimensional trait and temperament scales of the SNAP to the heterogeneous criteria provided by the PD scales of the SNAP and the diagnostic classes measured by the SCID.

The comparative validity results varied somewhat as a function of criterion type. For both SNAP scale sets, the results revealed somewhat better discriminant validity for the RC scales as well as greater overall concurrent validity. In addition, although both the standard and restructured clinical scales added incrementally to the prediction of the SNAP, the magnitude of the incremental effect of the RC scales in predicting the trait and temperament scales was 2.8 and 1.8 times larger than that of the standard clinical scales in the veteran and clinic samples, respectively. Likewise, the incremental effect of the RC scales in predicting SNAP PD scales was 2.3 and 1.4 times larger than that of the standard clinical scales in the veteran and clinic samples, respectively. Thus, the RC scales appear to represent an improvement in prediction of both discrete and complex criteria as measured by the SNAP. In addition, although the RC scales yielded a smaller proportion of large correlations with scales tapping negative affectivity/neuroticism than did the standard clinical scales, the validity patterns associated with both sets of clinical scales generally resembled that of the nonspecific RCd scale. Thus, although somewhat improved, the discriminant validity coefficients of the RC scales appear to be higher than would be ideal.

The SCID validity results reflected a somewhat different pattern. Here, few differences emerged in the overall and incremental prediction of the four primary diagnostic classes assessed by the SCID in this study. In fact, the only significant difference between the scale sets was for prediction of substance use disorders, with the RC scales adding incrementally to the prediction of both current and lifetime substance use disorders, and this effect was carried exclusively by RC4. No other incremental effects for either scale set were significant. It is important to note that it is likely that all of the SCID results were attenuated somewhat because of the use of binary diagnostic data, so replication using nonbinary syndromal data are an important next step to understand the true magnitude of these effects and whether any incremental advantage accrues across scale sets.

Discussion of Individual RC Scales

As promised by Tellegen et al. (2003), the restructuring of the clinical scales appears to have improved their distinctiveness while maintaining some conceptual connection to the original clinical scales. However, data presented here and by Tellegen et al. (2003) suggest that the RC scales cannot be interpreted simply through

traditional codetype analyses using accumulated evidence on the standard clinical scales. Thus, in this section, we focus on providing interpretive guidance for each RC scale on the basis of our validity data.

One of the most interesting RC scales is RCd, which was used both to help strip nonspecific variance from the other RC scales and to provide an overall index of distress akin to the traditional method of assessing profile elevation (e.g., Greene, 2000). As indexed by its moderate to strong correlations with the standard clinical scales, RCd does serve as a marker of profile distress. However, in addition, it has interesting correlates of its own. Within the RC scales, RCd was most strongly related to RC2 and RC7. When related to the SNAP, RCd correlated most strongly with many scales relevant to negative emotionality and low positive emotionality, although its relations were somewhat stronger with negative emotionality than positive emotionality in our data. Thus, RCd appears to have behaved, as Tellegen et al. (2003) intended, as a marker of the higher order pleasantness–unpleasantness dimension of affect (Tellegen et al., 1999; Watson & Tellegen, 1985).

A number of RC scales—RCd, RC2, and RC7—seem particularly useful for the assessment of depression and anxiety. Of the restructured scales, RCd was the strongest predictor of current and lifetime ratings of depressive disorders, which also were marked to a lesser extent by RC1, RC2, and RC7. Given the theoretical underpinnings of the RC scales, prominent dimensional models of affect (Tellegen et al., 1999; Watson & Tellegen, 1985) and of depressive and anxious symptomatology (Brown, Chorpita, & Barlow, 1998; Clark & Watson, 1991; Mineka, Watson, & Clark, 1998) provide an interesting framework within which to evaluate these results. In these models, anxiety and depressive disorders share a nonspecific component that includes both anxious and depressed mood as well as symptoms such as insomnia, irritability, and impaired concentration. In addition, depression is marked specifically by low positive affect, whereas anxiety disorders are marked by factors specific to each disorder, such as anxious arousal for panic disorder (Brown et al., 1998) or reexperiencing symptoms for posttraumatic stress disorder (Simms et al., 2002). As described earlier, Tellegen et al. (2003) constructed RC2 to represent low positive affect, RC7 to reflect negative affectivity, and RCd to represent general hedonic valence.

That RC2 predicts syndromal depression less strongly than RCd likely reflects both the anhedonic flavor of RC2 and the nonspecificity of many features of SCID-rated depression. Nevertheless, it seems that using RCd and RC2 in concert can be helpful in more clearly understanding the particular nature of a patient's current distress. This conclusion is broadly consistent with the results presented by Tellegen et al. (2003, p. 32), who concluded, "the combined elevations on RCd and RC2 provide information regarding depression that is comparable to what is available from Clinical Scale 2 scores." However, it may be possible that different configurations of RCd and RC2 have different diagnostic implications. If, for example, both RCd and RC2 are clinically elevated, anhedonic depression might be a reasonable diagnostic hypothesis, whereas RCd elevated alone might signal general distress or perhaps clinical levels of depression marked more by sadness and nonspecific symptoms such as concentration problems, sleep difficulties, and rumination. Future studies are needed to examine such configural relationships in the RC scales.

As would be expected, RC7, which includes content broadly related to individual differences in negative emotionality, was related nonspecifically to both depressive and anxiety disorders. However, one concern raised by our data are the high correlations—which approached .80 in both samples—between RCd and RC7. Correlations of this magnitude suggest substantial construct overlap between these scales and raise questions about the necessity of having two scales tapping the same nonspecific variance. In contrast to the various depressive disorders which all share the specific factor of low positive emotionality, anxiety disorders (as noted earlier) are marked by disorder-specific factors. However, generalized anxiety disorder (GAD) has no disorder-specific factor but is characterized solely by the general negative emotionality factor (Watson, 1999). Thus, it may be that in constructing a core component for Scale 7, which is traditionally associated with anxiety disorders, Tellegen et al. (2003) developed a scale tapping the core of GAD and thus simply recreated the general negative emotionality factor. Thus, it would be interesting to see whether additional scales could be developed from the MMPI–2 item pool to assess the specific components of other anxiety disorders.

The restructuring of Scales 1 and 3 resulted in one coherent scale (RC1) reflecting a broad range of somatic complaints as well as a second scale (RC3) whose interpretive meaning appears to be quite different from its predecessor. RC1 and Scale 1 equally predicted somatoform disorders, and both did so better than Scale 3. In addition, RC1 was broadly related to negative emotionality, self-harm behavior, interpersonal mistrust, and SCID-rated depression and anxiety. RC3, in contrast, was not strongly related to any Axis I diagnostic class examined in this study but rather appeared to be closely related to the interpersonal mistrust and persecutory ideation that characterize Paranoid PD. Moreover, RC3's pattern of relations with the SNAP and SCID was markedly different than those of the original Scale 3. Thus, the different conceptual meaning of RC3 cannot be overemphasized, and future studies are needed to understand and document its correlates and construct validity.

The restructuring of Scale 4 appears to have resulted in a scale that is far less imbued with nonspecific distress. Instead, RC4 was related more specifically to externalizing and disinhibitory psychopathology, such as substance use disorders, antisocial and borderline PDs, and traits of manipulativeness, aggression, and impulsivity. On a related note, RC9 yielded similar correlates to RC4 but actually was more strongly correlated with manipulativeness and aggression and also had moderate relations with exhibitionism (e.g., wanting to be the center of attention) and unusual perceptual experiences. At the diagnostic level, RC9 was related most strongly to narcissistic and histrionic PDs but also at a moderate level with antisocial and borderline PDs. Thus, although the empirical data related to the traditional 4–9/9–4 codetype cannot be assumed to apply to the RC scales, RC4 and RC9 together seem to represent, in perhaps a conceptually clearer way, many of the characteristics traditionally associated with individuals producing a 4–9/9–4 profile (e.g., Butcher & Williams, 2000; Greene, 2000).

RC6 and RC8 also yielded somewhat clearer relations than their standard clinical scale counterparts. Although moderately correlated, RC6 and RC8 yielded meaningful and somewhat differentiable correlational patterns with the SNAP. Specifically, RC6 correlated most strongly with a scale tapping interpersonal mistrust

and more moderately with a scale measuring unusual perceptual experiences, whereas an opposite pattern was observed for RC8, which correlated most strongly with unusual perceptual experiences and more moderately with interpersonal mistrust. At the diagnostic level, RC6 correlated most strongly with paranoid and schizotypal PDs, whereas RC8 was more specifically related to schizotypal PD. Our samples included too few individuals with diagnosable psychotic disorders to provide meaningful, definitive Axis I syndromal correlates for RC6 and RC8; future studies are needed to explore such relations.

Summary and Conclusions

Our study is not without limitations. First, we focus only on relations between the RC scales and their standard clinical scale counterparts. Nichols (2004) has argued that the RC scales represent merely a repackaging of other established MMPI–2 content or supplementary scales. Indeed, several RC scales in the present data sets are highly correlated with other scales routinely scored on the MMPI–2, largely because of shared items. For example, RC1 correlates .95 with Health Concerns; RC3 correlates .93 with the Cynicism content scale; and RCd correlates .95 with Welsh's Anxiety scale. However, this does not necessarily represent a threat to the validity of the RC scales, as Tellegen (2005) claims that the RC scales were never intended to be completely novel or unique scales. Nevertheless, studies examining the overlap between the RC scales and other MMPI–2 scale sets are needed to explicate this issue fully.

Likewise, more data are needed to examine the concurrent and predictive validity of the RC scales with respect to other established measures of personality and psychopathology as well as important criteria of interest (e.g., therapy process and outcome variables, collateral ratings of symptoms, etc.). Finally, characteristics of our samples—which were both samples of convenience—may limit generalizability to other samples. For example, the veteran sample consisted predominantly of Caucasian men, and all were originally from the state of Iowa; thus, whether these results are generalizable to female veterans or ethnically diverse veterans from different geographic regions of the country may need to be addressed in future studies.

In summary, the RC scales appear to represent a modest psychometric improvement over the standard clinical scales. Their greater internal coherence and discriminant validity have resulted in cleaner and conceptually clearer relations both within the MMPI–2 and with other measures of personality and psychopathology. In particular, the attempt to remove nonspecific distress variance appears to have been somewhat successful, which should result in clinical profiles that are more easily interpreted. Moreover, the RC scales appear to predict complex criteria such as psychiatric diagnoses at least as well as their standard scale counterparts, but replication of these findings is necessary with nonbinary diagnostic data. Also, despite the temptation to do so, it also is apparent that the RC scales cannot be interpreted on the basis of previous empirical studies of the original scales; the RC scales represent new measures whose meanings now must be determined empirically.

References

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

Arbisi, P. A. (2004, March). *Further research of the MMPI–2 restructured clinical scales.* Symposium conducted at the annual meeting of the Society for Personality Assessment, Miami, FL.

Brown, T. A., Chorpita, B. F., & Barlow, D. H. (1998). Structural relationships among dimensions of the *DSM–IV* anxiety and mood disorders and dimensions of negative affect, positive affect, and autonomic arousal. *Journal of Abnormal Psychology, 107,* 179–192.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI–2 (Minnesota Multiphasic Personality Inventory–2): Manual for administration and scoring.* Minneapolis, MN: University of Minnesota Press.

Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W. G., & Kaemmer, B. (2001). *MMPI–2 (Minnesota Multiphasic Personality Inventory–2): Manual for administration and scoring* (Rev. ed.). Minneapolis, MN: University of Minnesota Press.

Butcher, J. N., & Williams, C. L. (2000). *Essentials of MMPI–2 and MMPI—A interpretation.* Minneapolis, MN: University of Minnesota Press.

Clark, L. A. (1993). *Schedule for Nonadaptive and Adaptive Personality: Manual for administration, scoring, and interpretation.* Minneapolis, MN: University of Minnesota Press.

Clark, L. A., Simms, L. J., Wu, K. D., & Casillas, A. (in press). *Schedule for Nonadaptive and Adaptive Personality: Manual for administration, scoring, and interpretation* (2nd ed.). Minneapolis, MN: University of Minnesota Press.

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100,* 316–336.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Doebbeling, B. N., Jones, M. F., Hall, D. B., Clarke, W. R., Woolson, R. F., et al. (2002). Methodological issues in a population-based health survey of Gulf War veterans. *Journal of Clinical Epidemiology, 55,* 477–487.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1997). *Structured Clinical Interview for Axis I DSM–IV Disorders—Patient Edition (SCID—I/P).* New York: Biometrics Research Department, New York State Psychiatric Institute.

Graham, J. R. (2000). *MMPI–2: Assessing personality and psychopathology* (3rd ed.). New York: Oxford University Press.

Greene, R. L. (2000). *The MMPI–2: An interpretive manual* (2nd ed.). Boston: Allyn & Bacon.

Hathaway, S. R., & McKinley, J. C. (1943). *The Minnesota Multiphasic Personality Inventory.* Minneapolis, MN: University of Minnesota Press.

Helmes, E., & Reddon, J. R. (1993). A perspective on developments in assessing psychopathology: A critical review of the MMPI and MMPI–2. *Psychological Bulletin, 113,* 453–471.

Horn, J. L., Wanberg, K. W., & Appel, M. (1973). On the internal structure of the MMPI. *Multivariate Behavioral Research, 8,* 131–171.

Iowa Persian Gulf Study Group. (1997). Self-reported illness and health stats among Gulf War veterans. *JAMA, 277,* 238–245.

McFall, R. M., & Treat, T. A. (1999). Quantifying the information value of clinical assessments with signal detection theory. *Annual Review of Psychology, 50,* 215–241.

Mineka, S., Watson, D., & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology, 49,* 377–412.

Nichols, D. S. (2004). *Content scales with clinical roots: A review and critique of the MMPI–2 restructured clinical scales.* Unpublished manuscript.

Osberg, T. M., Haseley, E. N., Kamas, M. M., & Henderson, D. (2004,

March). *The MMPI–2 clinical scales and RC scales: Comparability of indices of profile elevation and general maladjustment in young adults.* Paper presented at the annual meeting of the Society for Personality Assessment, Miami, FL.

Pfohl, B., Blum, N., Zimmerman, M., & Stangl, D. (1989). *Structured interview for DSM–III–R* personality (SIDP-R). Iowa City: University of Iowa, Department of Psychiatry.

Reddon, J. R., Marceau, R., & Jackson, D. N. (1982). An application of singular value decomposition to the factor analysis of MMPI items. *Applied Psychological Measurement, 6,* 275–283.

Simms, L. J., Casillas, A., Clark, L. A., & Watson, D. (2004, March). *Psychometric evaluation of the restructured clinical scales of the MMPI–2.* Paper presented at the annual meeting of the Society for Personality Assessment, Miami, FL.

Simms, L. J., Watson, D., & Doebbeling, B. N. (2002). Confirmatory factor analyses of posttraumatic stress symptoms in deployed and nondeployed veterans of the Gulf War. *Journal of Abnormal Psychology, 111,* 637–647.

Tellegen, A. (2005, March). [Discussant]. In Y. S. Ben-Porath (Chair), *MMPI–2: Restructured clinical (RC) scales.* Symposium presented at the annual meeting of the Society for Personality Assessment, Chicago.

Tellegen, A., Ben-Porath, Y. S., McNulty, J. L., Arbisi, P. A., Graham, J. R., & Kaemmer, B. (2003). *The MMPI–2 restructured clinical (RC) scales: Development, validation, and interpretation.* Minneapolis, MN: University of Minnesota Press.

Tellegen, A., Watson, D., & Clark, L. A. (1999). On the dimensional and hierarchical structure of affect. *Psychological Science, 10,* 297–303.

Waller, N. G. (1999). Searching for structure in the MMPI. In S. E. Embretson & S. L. Hershberger (Eds.), *The new rules of measurement* (pp. 185–218). Mahwah, NJ: Erlbaum.

Watson, D. (1999). Dimensions underlying the anxiety disorders: A hierarchical perspective. *Current Opinion in Psychiatry, 12,* 181–186.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98,* 219–235.

Williams, J. B. W., Gibbon, M., First, M. B., Spitzer, R. L., Davies, M., et al. (1992). The structured clinical interview for *DSM–III–R* (SCID). Part II: Multisite test–retest reliability. *Archives of General Psychiatry, 49,* 630–636.

Zanarini, M. C., Skodol, A. E., Bender, D., Dolan, R., Sanislow, C., Schaefer, E., et al. (2000). The Collaborative Longitudinal Personality Disorders Study: Reliability of Axis I and II diagnoses. *Journal of Personality Disorders, 14,* 291–299.

Received August 17, 2004
Revision received April 6, 2005
Accepted April 12, 2005 ■