EXAMINING CONVERGENT AND DISCRIMINANT VALIDITY AND INCREMENTAL
UTILITY OF THE KOREAN MMPI-2 RESTRUCTURED CLINICAL (RC) AND CLINICAL
SCALES IN KOREAN PSYCHIATRIC AND NORMATIVE COUPLE SAMPLES

Amy S. Dykhouse

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Department of Psychology

Central Michigan University
Mount Pleasant, Michigan
March 2014

ACKNOWLEDGEMENTS

ABSTRACT

EXAMINING CONVERGENT AND DISCRIMINANT VALIDITY AND INCREMENTAL
UTILITY OF THE KOREAN MMPI-2 RESTRUCTURED CLINICAL (RC) AND CLINICAL
SCALES IN KOREAN PSYCHIATRIC AND NORMATIVE COUPLE SAMPLES

by Amy S. Dykhouse

The symptom substantive Clinical Scales of the Minnesota Multiphasic Personality

Inventory-2 (MMPI-2; Butcher, Graham, Ben-Porath, Tellegen, Dahlstrom, & Kaemmer, 1989)

went through a substantial revision process creating scales that (a) showed lower

intercorrelations among scales and (b) comprised more homogeneous item content. These new

scales are called the Restructured Clinical Scales (RC; Tellegen et al., 2003). Multiple studies

have provided evidence supporting their utility. Zero-order correlations of RC Scales and

Clinical Scales with both self-report and external criterion measures of psychopathology show

most RC Scales do provide similar to greater convergent and discriminant validity than that of

Clinical Scale counterparts. RC Scales also contribute unique information above that of

respective Clinical Scales in prediction of a range of psychological symptoms (Sellbom & Ben-

Porath, 2005; Simms, Casillas, Clark, Watson, & Doebbeling, 2005).

However, several existing issues point to the need for further research to be conducted.

Research has reported RC Scales fail to measure the same construct as do respective Clinical

Scales (Bolinskey & Nichols, 2011; Butcher, Hamilton, Rouse, & Cumella, 2006). In addition,

RC Scale ability to contribute in the prediction of psychological symptoms to that of Clinical

Scales has been established in simple criterion measures, such as anxiousness, but not in more

complex measures like diagnostic status (Nichols, 2006). Furthermore, Koreans are one of the

largest immigrant groups in the United States (United States Census, 2010), and with the recent

publication of the Korean version of the MMPI-2-RF (Han, Moon, Lee, & Kim, 2011), clinicians

will likely have interest in and need for information on the equivalence of RC Scales with that of the Clinical Scales of the Korean MMPI-2 (Kim et al., 2005) using Korean samples.

Due to the (a) disagreement over RC Scale assessment value, and (b) need for studies using Korean samples, additional research on the RC Scales is required to examine convergent and discriminant validity, as well as incremental utility. This study therefore examined discriminant and convergent validity and incremental utility of RC and Clinical Scale using several external criterion measures of psychological symptoms. The participants included a sample of 395 Korean psychiatric patients (176 men and 219 women) and a sample of 108 married couples ($N = 216$) from the Korean MMPI-2 normative sample. Criterion measures consisted of Patient Description Form (PDF; Graham, Ben-Porath, and McNulty, 1999) and diagnostic status (Structured Clinical Interview for Diagnosis on DSM-IV; SCID) for the psychiatric sample and the Spouse Rating Form (SRF; Butcher et al., 1989) for the couple sample. All patients completed the Korean version of the MMPI-2 (Kim et al., 2005), and therapists completed SCID diagnoses as well as a shortened version of the PDF. Each individual in the couple sample completed the Korean MMPI-2 and rated his or her spouse on the SRF.

In roughly half of SCID diagnostic categories, PDF Scales, and SRF Scales, RC Scales (a) yielded similar to greater convergent and discriminant validity compared to Clinical Scale versions and (b) contributed similar to greater proportion of incremental utility to that of respective Clinical Scales than that of the Clinical Scale to the RC Scale version. These findings add support for the comparative efficacy of this newly published and drastically shortened Korean MMPI-2-RF (Han et al., 2011) to its predecessor, the Korean MMPI-2 (Kim et al., 2005). Further research performed using more commonly used and thoroughly validated external criterion measures will enable comparability of results with a greater number of similar studies.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Dating back to the 1950s, the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1943) has a long history of being commonly and successfully used as an objective measure of personality and psychopathology in both applied and research settings (Ben-Porath, 2012). Items of the measure's eight Clinical Scales were selected via the empirical keying approach differentiating patients of eight different diagnostic groups (e.g., hypochondriasis, depression, hysteria, psychopathic deviance, paranoia, psychasthenia, schizophrenia, and hypomania) from non-patients. This method of scale construction resulted in scales being highly correlated to each other and containing heterogeneous item content.

In the revision process to the Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Butcher et al., 1989) very few of the Clinical Scales were altered for the purpose of maintaining continuity between test versions (Ben-Porath & Tellegen, 2008). As a result, these problems still existed. Authors therefore undertook the process of revising the Clinical Scales to construct scales that were less intercorrelated and contained more homogeneous item content (Ben-Porath, 2012). The resulting scales were the Restructured Clinical Scales (RC; Tellegen et al., 2003) included in the Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008).

The success of the Clinical Scales as being accurate assessments of conceptually related self-report and external criterion measures of psychological symptoms has been replicated in hundreds of studies (Graham, 2006).   Due to their effectiveness, clinicians heavily rely on Clinical Scales when interpreting test results for their clients. Research has found similar

support for the utility of the RC Scales in terms of good convergent and discriminant validity (Sellbom & Ben-Porath, 2005; Sellbom, Ben-Porath, et al., 2006; Simms et al., 2005).

In contrast, Bolinskey and Nichols (2011) have argued that RC Scales assess different constructs as do the Clinical Scales. An additional criticism of the RC Scales argued by Nichols (2006) is that comparisons of incremental utility between scale versions should be performed using complex, broad-band measures of psychopathology (e.g., SCID diagnoses) as criterion. This analysis will provide a more rigorous assessment of RC Scale ability to capture heterogeneous symptom constructs as well as do that of the Clinical Scales.

As of 2013, over 150 non-English translations have been published of the original MMPI and 21 of the MMPI-2 (University of Minnesota Press). Research has provided evidence scales of these international MMPI versions are able to reliably predict behavior and personality and to show construct equivalence across many cultural groups (Butcher, Lim, & Nezami, 1998). Despite the availability of an extensive number of translations of the MMPI-2, to the authors' knowledge, no studies have compared incremental utility of the RC to that of the Clinical Scales in non-U.S. samples. With the Korean MMPI-2-RF being one of only three non-English MMPI-2-RF versions available (University of Minnesota Press, 2013), clinicians have need of further information on the newly published RC Scales and their interpretation using Korean samples.

Due to the (a) controversy surrounding utility of the RC Scales and (b) relative non-existence of RC Scale research conducted using Korean samples, a thorough examination of the unique assessment information contributed by the RC in comparison to their previous Clinical Scale versions is essential. This study therefore proposed to examine convergent and discriminant validity through zero order correlations of RC and Clinical Scales with external criterion measures. These analyses will address the criticism that scale versions measure

different constructs (Bolinskey & Nichols, 2011). The study also investigated incremental utility of RC and Clinical Scales to provide evidence of RC Scales to contribute valuable information to that of the Clinical Scales in prediction of complex external criterion measures of psychological symptoms. Participants include both a Korean psychiatric hospital sample and a Korean normative sample consisting of married couples.   External criterion measures consisted of Structured Clinical Interview for Diagnosis on DSM-IV (SCID) diagnosis, Patient Description Form (PDF), and Spouse Rating Form (SRF).

CHAPTER II

LITERATURE REVIEW

MMPI-2 Restructured Clinical (RC) Scales

Development of RC Scales

To accomplish the task of revising the Clinical Scales to form the RC Scales, Tellegen et al. (2003) used both exploratory factor and correlational analyses. The first step performed was to separate a construct of demoralization from the MMPI-2 item set in order to reduce the high intercorrelations among the Clinical Scales. This demoralization factor had been noted to be a common factor across various forms of psychopathology and had been included in many psychiatric screening measures. Most likely the high saturation of this demoralization factor was due to the procedures used in creating the original Clinical Scales. Items were selected that differentiated psychiatric patients of select psychological disorders from those of a nonclinical control group. As a result, some of the items were markers of a specific disorder; however, some items differentiated clinical from non-clinical symptoms. Authors selected items for this demoralization factor from Clinical Scales of Depression (D) and Psychasthenia (Pt). These two scales most strongly corresponded to measures of Positive and Negative Affect, of which research had shown to explain a large proportion of common variance in self-report measures (Ben-Porath, 2012).

The next step of the process was to subject the eight Clinical Scales (plus Masculine-Feminine [Mf] and Social Introversion [Si]) to factor analyses. Item markers of Demoralization were included in these separate analyses to identify and remove additional items related to Demoralization from each Clinical Scale. From these analyses, seed scales for Demoralization

and each of the Clinical Scales were formed. Authors intended for this process to improve discriminant validity and item homogeneity of each scale.

The last step consisted of performing correlational analyses on the 12 seed scales using the full 567 MMPI-2 item set. An item was added to the seed scale if its correlation was stronger with the seed scale than with the other 11 scales, or if the correlation coefficient was sufficiently large. Some scales necessitated higher or lower criteria of correlation to be met for inclusion in the scale (see Tellegen et al., 2003 for description). The final item set of the RC Scales was reported to be shorter, more reliable, show weaker intercorrelations with other scales, contain fewer items assessing Demoralization, and show improved convergent and discriminant validity (Ben-Porath, 2012). See Table 1 for description of item content and item count for Clinical (top panel) and RC Scales (bottom panel).

Criticism of RC Scales

One of the main issues raised regarding utility of the RC Scales is that they do not measure the same constructs as do respective Clinical Scales. For instance, Butcher, Hamilton, Rouse, and Cumella (2006) and Nichols (2006) point out the negative correlation of Cynicism (RC3) with its Clinical Scale counterpart Hysteria (Hy), most likely an artifact of the decision of RC scale developers to reverse keying so that scale elevations reflect endorsement rather than denial of cynicism. Nichols (2006) has used the term construct drift to label this difference in symptoms assessed between scale versions. He argues the constructs of the original Clinical Scales were altered when forming the RC Scale counterpart. Content not present in the original Clinical Scale was added to the RC version during the step of the revision process in which all 567 items from the MMPI-2 were correlated with seed scales. The impact of adding these new

items is severe in some scales, because the proportion of items added in this step outnumbered

that of the original seed scales by an average ratio of 3 to 1. For example, of the items

Table 1. *Clinical and RC Scales: MMPI-2 Scale Names, Item Count, and Content Descriptions*

Scale	Name (Item count)	Description
Clinical Scale		
Hs	Hypochondriasis (32)	Variety of somatic complaints
D	Depression (57)	Lack of hope, life dissatisfaction, poor morale
Hy	Hysteria (60)	Immature, narcissistic, somatic complaint
Pd	Psychopathic Deviate (50)	Difficulty with authority, family problems
Pa	Paranoia (40)	Suspiciousness, delusions of persecution
Pt	Psychasthenia (48)	Excessive doubts, compulsions, obsessions
Sc	Schizophrenia (78)	Thought, mood, and behavior disturbances
Ma	Hypomania (46)	Elevated mood, accelerated speech or activity
Restructured Clinical (RC) Scale		
RCd	Demoralization (24)	General unhappiness
RC1	Somatic Complaints (27)	Wide range of physical health complaints
RC2	Low Positive Emotions (17)	Deficit of positive emotional responses
RC3	Cynicism (15)	General thoughts of distrust of others
RC4	Antisocial Behavior (22)	Failure to follow laws or rules
RC6	Ideas of Persecution (17)	Feelings of threat from others regarding the self
RC7	Dysfunctional Negative Emotions (24)	Feelings of anxiety or anger that are unhealthy
RC8	Aberrant Experiences (18)	Unusual thoughts or perceptions
RC9	Hypomanic Activation (28)	Extreme feelings of aggression, impulsivity

Note. Adapted from *MMPI-2 Assessing Personality and Psychopathology, 4th Edition* (Graham, 2006) and *MMPI-2-RF Manual for Administration, Scoring, and Interpretation* (Ben-Porath & Tellegen, 2008).

comprising Dysfunctional Negative Emotions (RC7) and Hypomanic Activation (RC9), 67% and

71%, respectively, are not contained in the original scale.

Recently, Bolinskey and Nichols (2011) examined the lack of construct equivalence

between RC4, RC7, and RC9 and respective Clinical Scale versions using principal components

analyses with Varimax rotation. Results showed that a greater proportion of items of RC Scales

loaded on a general maladjustment factor than did respective Clinical Scales. These results

provided evidence that the process chosen by developers of the RC Scales was not successful in

removing non-specific content related to demoralization; proportion of items containing demoralization content in RC Scales actually increased from that of the Clinical Scale versions.

In addition, Nichols (2006) asserts that research presented by developers of the RC Scales that supported the predictive utility of RC Scales above that of Clinical Scales presented a biased representation due to choice of external criterion measures. Due to differences in method of Clinical and RC Scale construction (empirical keying and factor analysis, respectively), RC Scales will provide greater utility in prediction of relatively narrow symptom constructs, such as anxiousness or somatic complaints, whereas Clinical Scales will contribute greater utility in prediction of more complex criterion (e.g., clinical diagnoses). Therefore, Nichols argues that incremental utility of the RC Scales above that of the Clinical Scales has yet to be fully assessed.

<div align="center">Clinical Scales vs. RC Scales</div>

Convergent and Discriminant Validity

In spite of these criticisms, research using a variety of criterion measures has provided evidence showing improved utility for a majority of the RC Scales over that of Clinical Scale versions. Proponents argue RC, as compared to Clinical Scales, show stronger convergent and weaker discriminant correlations with criterion measures due to the methods used in the process of revision: RC Scales were formed to be consistent in structure and substance while being distinctive from each other (Sellbom & Ben-Porath, 2005). For example, Sellbom and Ben-Porath compared RC and Clinical Scale convergent validity using the Multidimensional Personality Questionnaire (MPQ; Patrick, Curtin, & Tellegen, 2002), a self-report measure consisting of 11 scales, of which ten form three higher-order dimensions (positive emotionality, negative emotionality, and constraint). Participants comprised 811 college students (290 men

and 521 women). Results of zero-order comparisons revealed most MPQ scales showed stronger correlations with conceptually relevant RC Scales than with respective Clinical Scales. Examples included Stress Reaction with RCd (r = .68) than D (r = .52), Well-Being with RC2 (r = -.72) than with D (rs = -.63), Alienation with RC3 (r = .54) than with Hy (r = .08), Control with RC4 (r = -.36) than with Pd (r = -.27), Alienation with RC6 (r = .62) than with Pa (r = .49), Absorption with RC8 (r = .52) than with Sc (r = .37), and Aggression with RC9 (r = .57) than with Ma (r = .42). In contrast, correlations of Stress Reaction with RC1 were similar (r = .50) to those with Hs (r = .53), and both Well-Being ([D] r = -.63 vs. [RCd] r = -.59) and Stress Reaction ([Pt] r = .75 vs. [RC7] r = .71) showed slightly stronger relationships with Clinical than with RC Scale counterparts.

A majority of RC Scales also showed greater discriminant validity of conceptually relevant MPQ Scales than Clinical Scale counterparts. Discriminant validity (r_{dis}) was assessed by subtracting the (a) average correlation with all conceptually unrelated criterion measures (r_{mean}) from the (b) correlation with conceptually relevant criterion measure (r). Larger (positive) differences between correlation coefficients indicated better discriminant validity. For example RCd showed greater discriminant validity of Stress Reaction (r_{dis} = [r = .68] − [r_{mean} = .28] = .40) than did D (r_{dis} = [r = .52] − [r_{mean} = .22] = .30). Other examples in which RC Scales provided greater discriminant validity than respective Clinical Scales included RC2 of Well Being (r_{dis} = .48) than D (r_{dis} = .42), RC3 of Alienation (r_{dis} = .36) than Hy (r_{dis} = -.03), RC4 of Control (r_{dis} = .12) than Pd (r_{dis} = -.02), RC6 of Alienation (r_{dis} = .44) than Pa (r_{dis} = .28), RC8 of Absorption (r_{dis} = .34) than Sc (r_{dis} = .06), and RC9 of Aggression (r_{dis} = .33) than Ma (r_{dis} = .18). Several RC and respective Clinical Scales showed similar discriminant validity of MPQ Scales including Stress Reaction ([Hs] r_{dis} = .35 vs. [RC1] r_{dis} = .33) and Stress Reaction ([Pt] r_{dis} = .48

8

vs. [RC7] $r_{dis} = .46$). D, however, showed greater discriminant validity of Well Being ($r_{dis} = .42$) than did RCd ($r_{dis} = .31$).

Sellbom, Graham, and Schenk (2006) also assessed convergent and discriminant validity of the RC Scales by correlating them with factor scales derived from factor analyses of the Multiaxial Diagnostic Inventory (MDI; Doverspike, 1990), a self-report measure of psychological symptoms. The sample comprised private practice clients ($N = 647$; 240 men and 407 women) with most common diagnoses of Depressive (42%), Anxiety (34%), and Adjustment Disorders (33%). Analyses performed on the same sample assessing Clinical Scale utility were published in Sellbom, Graham, and Schenk (2005), which allowed for comparison of utility among RC and respective Clinical Scales. Most MDI factor scales showed stronger correlations with conceptually relevant RC Scales than with respective Clinical Scale versions. Examples included Somatization with RC1 ($r = .69$) than with Hs ($r = .62$), Paranoia/Mistrust with RC3 ($r = .52$) than with Hy ($r = .06$), Antisocial with RC4 ($r = .61$) than with Pd ($r = .36$), Paranoia/Mistrust with RC6 ($r = .53$) than with Pa ($r = .41$), Bizarre Experiences with RC8 ($r = .57$) than with Sc ($r = .51$), and Mania with RC9 ($r = .49$) than with Ma ($r = .37$). Depression showed similar strength correlations with RCd ($r = .65$) and D ($r = .67$). There were exceptions in which Clinical Scales showed stronger correlations with conceptually relevant MDI factor scales than did respective RC Scales: Depression with D ($r = .67$) than with RC2 ($r = .53$); Panic/Anxiety with Pt ($r = .54$) than with RC7 ($r = .48$).

A majority of RC Scales showed greater discriminant validity than did respective Clinical Scales: Paranoia/Mistrust with RC3 ($r_{dis} = .25$) than with Hy ($r_{dis} = -.26$), RC4 of Antisocial ($r_{dis} = .32$) than with Pd ($r_{dis} = .07$), RC6 of Paranoia/Mistrust ($r_{dis} = .23$) than with Pa ($r_{dis} = .07$), RC8 of Bizarre Experiences ($r_{dis} = .20$) than with Sc ($r_{dis} = .08$), and RC9 of Mania ($r_{dis} = .24$) than

with Ma (r_{dis} = .14). In comparison, both RCd and RC2 showed weaker discriminant validity of Depression (rs_{dis}= .23 and .28, respectively) than did D (r_{dis} = .34). RC1 also showed weaker discriminant validity of Somatization (r_{dis} = .28) than with Hs (r_{dis} = .34), as did RC7 of Panic Anxiety (r_{dis} = .03) than with Pt (r_{dis} = .14).

Along with studies using self-report measures as criterion, research of Sellbom, Ben-Porath, et al. (2006) demonstrated most RC Scales do provide similar to greater convergent and discriminant validity than respective Clinical Scales when external criterion measures were used. Researchers examined correlates of RC Scales with conceptually related criterion measures in a university psychological clinic setting (228 men and 522 women). External criterion measures included Intake Form (Graham, Ben-Porath, & McNulty, 1999), Client Description Form (CDF; Graham et al., 1999), and Symptom Checklist-90-Revised-Analogue (SCL-90-R; Derogatis, Rickels, & Rock, 1976). RC Scales, with few exceptions, exhibited equivalent or greater convergent validity of hypothetically related criterion measures as that of respective Clinical Scales. RC3 (r = .14) and RC6 (r = .12) showed stronger correlation with CDF Scale Suspicious than did either Hy (r = -.05) or Pa (r = .06). RC4 also showed stronger correlation with CDF Scale Antisocial (r = .35) and slightly stronger correlation with Aggressive (r = .22) than did Pd (rs =.25 and .17, respectively). RC9, compared to Ma, showed similar to stronger correlations with SCL-90-R Hostility (rs =.16 vs. .12), CDF Antisocial (rs = .27 vs. .21), CDF Aggressive (rs = .20 vs. .14), and CDF Narcissistic (rs = .17 vs. .15). RC1 and Hs showed similar convergent validity of SCL-90-R Analogue Somatization (both rs =.28) and CDF Somatic Symptoms (both rs =.25).

Convergent validity of RC and Clinical Scales showed different patterns of utility across several of the conceptually relevant criterion measures. For example RC2 and D showed similar

strength correlations with Client Intake Depression ($rs = .28$ and $.29$, respectively) and SCL-90-R Interpersonal Sensitivity (both $rs = .24$), while SCL-90-R Depression and CDF Depressed were more strongly correlated with D ($rs = .38$ and $.41$, respectively) than with RC2 ($r = .27$ and $.33$, respectively). Results were similar when comparing convergent validity of RCd with that of D: both showed similar convergent validity of Client Intake Depression ($rs = .29$) and SCL-90-R Interpersonal Sensitivity ($rs = .24$), while D showed slightly greater convergent validity of SCL-90-R Depression ($r = .38$) and greater convergent validity of CDF Depressed ($r = .41$) as compared to that of RCd ($r = .33$ and $.35$, respectively). In addition, RC7 and Pt showed similar convergent validity of both Client Intake Anxiety ($rs = .15$ and $.13$, respectively) and SCL-90-R Anxiety ($rs = .21$ and $.22$, respectively); however, CDF Scale Anxious showed stronger correlations with Pt ($rs = .25$ and $.19$, respectively) and slightly stronger correlations with Obsessive-Compulsive ($rs = .18$ and $.14$, respectively) than with RC7. Convergent validity results comparing RC8 and Sc were not available due to low endorsement rate of symptoms of psychoticism within the sample.

For the most part, RC Scales showed equivalent or greater discriminant validity than respective Clinical Scales. RC3 showed better discriminant validity of CDF Suspicious ($r_{dis} = .09$) than did Hy ($r_{dis} = -.18$), as did RC4 of CDF Scales Antisocial ($r_{dis} = .23$) and Aggressive ($r_{dis} = .09$) than did Pd ($rs_{dis} = .12$ and $.03$, respectively). RC6 ($r_{dis} = .05$) also showed greater discriminant validity of CDF Suspicious than Pa ($r_{dis} = -.07$), and RC9 showed similar to slightly greater discriminant validity of CDF Antisocial ($r_{dis} = .18$), CDF Aggressive ($r_{dis} = .10$), CDF Narcissistic ($r_{dis} = .07$), and SCL-90-R Analogue Hostility ($r_{dis} = .12$) than did Ma ($rs_{dis} = .13$, $.05$, $.06$, and $.06$, respectively). RC1 and Hs showed similar discriminant validity of SCL-90-R Somatization ($rs_{dis} = .14$ and $.13$, respectively) and CDF Somatic Symptoms ($rs_{dis} = .17$ and $.15$,

respectively). As in comparisons of convergent validity, RCd, RC2, and RC7 exhibited varying utility compared to that of respective Clinical Scales across criterion measures. RCd showed similar to slightly weaker discriminant validity of Client Intake Depression (r_{dis} = .20), SCL-90-R Depression (r_{dis} = .17), and CDF Depressed (r_{dis} = .23) than did D (rs_{dis} = .21, .20, and .27, respectively). RC2 demonstrated greater discriminant validity of SCL-90-R Interpersonal Sensitivity (r_{dis} = .11) than did D (r_{dis} = .03) and similar discriminant validity of Client Intake Depression at (rs_{dis} = .22 and .21, respectively). D, however, did show slightly greater discriminant validity of SCL-90-R Depression (r_{dis} = .20) and greater discriminant validity of CDF Scale Depressed (r_{dis} = .27) than did RC2 (rs_{dis} = .15 and .21, respectively). RC7 also showed similar to slightly greater discriminant validity of Client Intake Anxiety (r_{dis} = .08), SCL-90-R Analogue Anxiety (r_{dis} = .05), and CDF Obsessive-compulsive (r_{dis} = .04) than did Pt (rs_{dis} = .03, .04, and .05, respectively), and slightly weaker discriminant validity of CDF Anxious (r_{dis} = .09 and .13, respectively).

Results of one of the few published studies using SCID diagnostic categories as external criterion measure (Simms et al., 2005) has yielded relatively moderate support of RC Scale utility compared to respective Clinical Scales. Part of these tempered results may be explained by the fact that relevant diagnoses necessary to compare utility of four RC Scales commonly yielding improved convergent and discriminant validity than their Clinical Scale counterparts (RC3, RC6, RC8, and RC9) were not available due to limited sample. SCID diagnoses were completed on a sample of military veterans from the first Gulf War (N = 567). Veterans were classified with the disorder as either present or absent within the following categories: Major Depression, Somatoform Disorder, Substance Abuse, and Anxiety Disorder. RC4 (r =.25) did show a stronger correlation with diagnosis of substance abuse than with Pd (r = .14); however,

most other RC Scales showed relatively similar to weaker correlations than with Clinical Scale counterparts with conceptually relevant diagnosis. For example, current diagnosis of depressive disorder showed a slightly stronger correlation with RCd ($r = .48$) than with D ($r = .44$), and current diagnosis of Somatoform Disorder showed similar correlation with RC1 ($r = .29$) than with Hs ($r = .28$). On the other hand, Clinical Scales showed slightly greater to greater convergent validity than did RC Scales of current diagnosis of Anxiety Disorder ([Pt] $r = .37$ vs. [RC7] $r = .32$) and of current diagnosis of Depressive Disorder ([D] $r = .44$ vs. [RC2] $r = .37$).

Comparisons of discriminant validity revealed RC Scales showed similar to weaker validity than respective Clinical Scales of all SCID diagnoses. For example, RCd ($r_{dis} = .27$) and D ($r_{dis} = .24$) showed similar discriminant validity of current diagnosis of Depressive Disorder, as did RC7 and Pt ($rs_{dis} = .13$) of current diagnosis of Anxiety Disorder. Both RC1 ($r_{dis} = .02$) Hs ($r_{dis} = .00$) showed similar discriminant validity of diagnosis of Somatoform Disorder (current). Slightly greater discriminant validity of current diagnosis of Depressive Disorder was demonstrated by D ($r_{dis} = .24$) than by RC2 ($r_{dis} = .19$).

Incremental Utility

Along with examination of convergent and discriminant validity, studies have also compared the amount of incremental utility contributed by either RC or respective Clinical Scales above that of the other in prediction of symptoms of psychopathology. Despite the fact that RC Scales were designed to assess one core element instead of the complete, multi-faceted character of respective Clinical Scales, researchers argue RC Scales do still predict a large proportion of variance in symptom prediction contained by the Clinical Scales, because they were designed to capture the key factor of each Clinical Scale (Tellegen, Ben-Porath, & Sellbom, 2009). Sellbom and Ben-Porath (2005), for instance, performed regression analyses to compare

the proportion of predictive utility contributed by the full set of Clinical Scales compared to that of the full set of RC Scales in prediction of each of the 11 self-report Multidimensional Personality Questionnaire Scales (MPQ; Patrick, Curtin, & Tellegen, 2002) and the three higher order scales in a college student sample (see earlier citation for sample description). The nine Clinical Scales (excluding Social Introversion [Si]) were entered together in each of the 14 models, and all 10 RC Scales were entered simultaneously in another set of 14 models. With the exception of models predicting Social Closeness ([RC] $R^2 = .20$ vs. [Clinical Scales] $R^2 = .26$) and Social Potency ([RC] $R^2 = .28$ vs. [Clinical Scales] $R^2 = .33$), proportion of predictive utility contributed by RC Scales was greater than that of the Clinical Scales in prediction of MPQ scales. Examples with more notable differences in predictive utility between scale groups included prediction of Aggression ([RC] $R^2 = .46$ vs. [Clinical Scales] $R^2 = .34$), Negative Emotionality ([RC] $R^2 = .66$ vs. [Clinical Scales] $R^2 = .55$), and Traditionalism ([RC] $R^2 = .22$ vs. [Clinical Scales] $R^2 = .13$). Averaged across models, RC Scales contributed greater proportion of predictive utility ($R^2_{mean} = .37$) than did Clinical Scales ($R^2_{mean} = .31$).

As compared to analyses comparing incremental predictive utility of RC versus that of the Clinical Scales as a group, Sellbom, Graham, et al. (2006) demonstrated individual RC Scales generally contribute significant proportions of incremental utility in prediction of self-reported psychological symptoms to that of their respective Clinical Scale (see earlier citation for sample details). Authors conducted a series of hierarchical regression analyses in which an RC Scale was entered in step one, and the corresponding Clinical Scale version was entered in step two, followed by reversing the order of entry. Results showed, on average, RC Scales contributed larger proportions of incremental utility to that of the Clinical Scales than that of the reverse order of entry in prediction of MDI factor scales ($\Delta R^2_{mean} = .12$ and .04, respectively).

For example, RC1 contributed greater proportion of incremental utility to Hs than that of Hs to

RC1 in prediction of Somatization (ΔR^2 = .09 and .00, respectively). Similar results were noted

for Paranoia/Mistrust ([RC3] ΔR^2 = .28 vs. [Hy] ΔR^2 = .01), Antisocial ([RC4] ΔR^2 = .25 vs. [Pd]

ΔR^2 =.00), Paranoia ([RC6] ΔR^2 = .13 vs. [Pa] ΔR^2 = .02, respectively), Bizarre Experiences

([RC8] ΔR^2 = .10 vs. [Sc] ΔR^2 = .04), and Mania ([RC9] ΔR^2 = .11 vs. [Ma] ΔR^2 =.01). There

were several exceptions. In prediction of MDI factor scale Depression, RCd and RC2

contributed smaller proportions of incremental utility to D (ΔR^2 = .04 and .00, respectively) than

that of the reverse order of entry (ΔR^2 = .06 and .17, respectively). In addition, the proportion of

incremental utility contributed by RC7 to Pt in prediction of Panic/Anxiety was less than that of

Pt to RC7 (ΔR^2 = .04 and .09, respectively).

Due to the fact that both RC and MDI factor scales were created through methods of

factor analyses and likely comprise similar levels of symptom homogeneity, researchers

performed additional hierarchical regression analyses using MDI Diagnostic Scales as criterion.

Authors reasoned that these analyses would provide a comparison in which the more symptom

heterogeneous Clinical Scales, instead of the RC Scales, might have the advantage in predictive

utility. Results showed RC Scales, on average, still contributed a greater proportion of

incremental utility to that of the Clinical Scales than that of the reverse order of entry (ΔR^2_{mean} =

.09 and .04, respectively). As in prediction of MDI factor scales, RC1, RC4, RC6, RC8, and RC9

contributed greater and both RCd and RC2 contributed smaller proportions of incremental utility

to that of respective Clinical Scales than that of the reverse order of entry in prediction of

conceptually relevant MDI Diagnostic Scales. Specific description of these increments are as

follows: RC1 to Hs than Hs to RC1 in prediction of Somatoform Disorder (ΔR^2 = .08 and .01,

respectively), RC3 to Hy (ΔR^2 = .07) than Hy to RC3 (ΔR^2 = .01) in prediction of Histrionic

Personality Disorder, RC4 to Pd ($\Delta R^2 = .25$) than Pd to RC4 in prediction of Antisocial Personality Disorder ($\Delta R^2 = .00$), RC6 to Pa ($\Delta R^2 = .10$) in prediction of Paranoid Personality Disorder than Pa to RC6 ($\Delta R^2 = .02$), RC8 to Sc ($\Delta R^2 = .09$) in prediction of Psychotic Disorder than Sc to RC8 ($\Delta R^2 = .00$), and RC9 to Ma ($\Delta R^2 = .11$) in prediction of Bipolar Disorder than Ma to RC9 ($\Delta R^2 = .01$). RCd and RC2 contributed smaller proportions of incremental utility over that of D in prediction of MDI Major Depressive Disorder ($\Delta R^2 = .04$ and $.00$, respectively) as D to RCd or RC2 ($\Delta R^2 = .06$ and $.17$, respectively). In contrast to results predicting MDI factor scales, RC7 contributed relatively similar and slightly greater proportion of incremental utility to Pt than Pt to RC7 in prediction of MDI Generalized Anxiety Disorder ($\Delta R^2 = .09$ and $.07$, respectively).

Simms et al. (2005) further examined incremental utility of the RC and Clinical Scales in prediction of both self-report and external criterion measures using an outpatient clinical sample and a military veteran sample. The first sample consisted of outpatient clients from a university psychology clinic ($n = 285$) with most common Axis I disorders of mood (23.9%), adjustment (7.1%), and anxiety disorders (6.8%). A military veteran sample from the first Gulf War ($n = 567$) comprised the second sample (see earlier citation for sample description). SCID diagnosis for current and lifetime psychiatric disorders information was available for the military veteran sample only. Self-report measure consisted of the Schedule for Nonadaptive and Adaptive Personality (SNAP; Clark, 1993). This measure contains 12 trait scales created through methods of factor analysis, as well as diagnostic scales designed to assess personality disorders listed in the Diagnostic and Statistical Manual of Mental Disorders (DSM–III–R; American Psychiatric Association, 1987). Researchers performed hierarchical regression and hierarchical logistic regression analyses in prediction of SNAP Scales and SCID diagnoses, respectively. Eight

Clinical Scales (Scales 1-9 with the exception of Masculinity-Femininity [Mf]) were entered as a block in step one, and the corresponding eight RC Scales (RC1, RC2, RC3, RC4, RC6, RC7, RC8, and RC9) were entered in step two, followed by reversing the entry order. Results provided evidence that both Clinical and the RC Scales contributed incremental utility in prediction of both SNAP Trait and Temperament Scales. In the military veteran and psychology clinic samples, RC Scales contributed greater proportion of incremental validity to Clinical Scales ($\Delta R^2_{mean} = .10$ and .09, respectively) than did Clinical to RC Scales (both $\Delta R^2_{mean} = .03$). In the military sample, examples included the following models predicting SNAP Trait Scales: Mistrust (RC: $\Delta R^2 = .15$ vs. Clinical Scale: $\Delta R^2 = .01$), Positive Temperament (RC: $\Delta R^2 = .11$ vs. Clinical Scale: $\Delta R^2 = .02$), and Disinhibition (RC: $\Delta R^2 = .12$ vs. Clinical Scale: $\Delta R^2 = .01$). RC and Clinical Scales contributed similar proportions of incremental utility in prediction of Dependency (both $\Delta R^2 = .03$) and in prediction of Detachment (RC: $\Delta R^2 = .09$ vs. Clinical Scale: $\Delta R^2 = .07$).

Regardless of the possible advantage for the Clinical Scales over that of the RC Scales resulting from use of more complex symptom traits such as personality disorders as criterion, results were comparable in the models predicting SNAP personality disorder scales. RC Scales contributed greater proportion of incremental utility to the Clinical Scales (both $\Delta R^2_{mean} = .09$) than Clinical to RC Scales ($\Delta R^2_{mean} = .03$ and .04) for military and psychology clinic samples, respectively. In the psychology clinic sample, examples included prediction of Antisocial Personality Disorder (PD) (RC: $\Delta R^2 = .20$ vs. Clinical Scale: $\Delta R^2 = .01$), Histrionic PD (RC: $\Delta R^2 = .15$ vs. Clinical Scale: $\Delta R^2 = .03$), Paranoid PD (RC: $\Delta R^2 = .09$ vs. Clinical Scale: $\Delta R^2 = .04$), and Narcissistic PD (RC: $\Delta R^2 = .11$ vs. Clinical Scale: $\Delta R^2 = .06$). RC and Clinical Scales

contributed similar proportions of incremental utility to each other in prediction of Schizotypal PD (both $\Delta R^2 = .04$).

In prediction of SCID diagnostic categories, results of hierarchical logistic regression analyses showed limited support for incremental predictive utility of either RC or Clinical Scales above that of other. In prediction of most diagnoses, neither RC nor Clinical Scales contributed significant proportion of incremental utility to the other. RC Scales did contribute a slightly greater proportion of incremental utility to the Clinical Scales than that of the Clinical Scales to the RC Scales in prediction of lifetime diagnosis of Depressive Disorders ($\Delta R^2 = .025$ and $\Delta R^2 = .014$, respectively), Anxiety Disorder ($\Delta R^2 = .011$ and $\Delta R^2 = .008$, respectively), and Somatoform Disorder ($\Delta R^2 = .012$ and $\Delta R^2 = .008$, respectively). In the model predicting lifetime diagnosis of substance abuse disorder, RC Scales did contribute significant and greater proportion of incremental utility to the Clinical Scales than that of the Clinical Scales to the RC Scales ($\Delta R^2 = .105$ and $.007$, respectively).

Researchers claimed use of binary criterion measure weakened results for both scale groups possibly causing neither RC nor Clinical Scales to contribute a significant proportion of incremental utility above that of the other in prediction of most SCID diagnoses. Authors argue, however, that the fact that RC Scales contributed relatively equivalent (though mainly non-significant) proportions of incremental utility to the Clinical Scales than that of the Clinical Scales to the RC Scales while containing fewer items was of value. In addition, applicable SCID diagnoses related to RC3, RC6, RC8, and RC9 were not presented due to limited sample. Results may have indicated more convincing support for either RC or Clinical Scale group above that of the other had additional diagnoses been available; however further analyses would be needed to confirm this.

Rationale of Current Study

The relatively recent revision of the Clinical to the RC Scales has drawn out several criticisms, one being the failure to measure similar constructs as their Clinical Scale counterparts (Butcher et al., 2006; Nichols, 2006). Although past publications have reported results supporting the cross-cultural equivalence of U.S. and Korean MMPI-2 versions (e.g., Han, 1996; Kim et al., 2005), studies have indicated minor disparities in how clients from Korean as compared to Western cultures report symptoms related to mental health. For instance Koreans more frequently endorse items related to high energy and activity than do people from Western cultures when instructed to appear well adjusted (Hahn, 2005). In addition, a common tendency of Korean patients is the reporting of a range of somatic complaints in place of expressing frustration or anger, a concept known as *Hwabyung* (Pang, 1990). Due to the possibility of disparities in symptom constructs assessed between RC and Clinical Scale versions, it is especially important that new research be performed to ensure interpretation of symptom endorsement remains pertinent for Korean psychotherapy clients.

Although several studies have demonstrated evidence that most RC Scales do provide similar to greater incremental utility as that of their Clinical Scale versions (e.g., Sellbom & Ben-Porath, 2005; Sellbom, Graham, et al., 2006), Nichols (2006) has advised that incremental utility research be performed using external criterion measures assessing a broad range of symptoms, such as SCID diagnoses, to provide more accurate comparison of assessment utility between RC and respective Clinical Scales. To the authors' knowledge, only one study has used SCID diagnosis as criterion (Simms et al., 2005), of which the sample comprised only a limited number of diagnostic categories. Due to sample limitations, as well as scarcity of research of this type in general, these findings are far from conclusive. For these reasons, this study will compare

19

convergent and discriminant validity, as well as incremental utility of the RC with that of the Clinical Scales to contribute to the limited research available addressing criticisms of lack of (a) construct equivalence and (b) accurate assessment of incremental predictive utility using complex and heterogeneous criterion measures. Analyses will be performed using multiple external criterion measures assessing both broad (e.g., SCID diagnoses) and specific symptom characteristics (e.g., Patient Description Form [PDF] and Spouse Rating Form [SRF] Scales) in Korean clinical and normative samples.

Criterion Measures

Author of the current study adopted methods similar to those of Barthlow, Graham, Ben-Porath, and McNulty (1999) and Sellbom, Graham, et al. (2006) in which a specific Clinical Scale and its respective RC Scale were chosen as predictors of rationally selected SCID diagnoses, PDF Scales, and SRF Scales. As Barthlow et al. suggests this method focuses not on achieving greatest prediction of each criterion measure, but on comparing two scales consisting of items of similar content and designed to measure similar symptoms in prediction of a specific symptom or behavior. One additional advantage of this method is that it reduces the chances of Type I errors by limiting criterion to those that have been established as relevant predictors by previous research. Selection of specific Clinical and RC Scales as predictors of diagnostic status were based on scale descriptions in the U.S. MMPI-2-RF manual (Ben-Porath & Tellegen, 2008) and previous research (Sellbom, Bagby, Kushner, Quilty, & Ayearst, 2012; Sellbom, Graham, et al., 2006). For example, Sellbom, Graham, et al. (2006) hypothesized the following RC and respective Clinical Scale versions as predictors of diagnostic status: RC1 and Hs of Somatoform Disorder; RCd, RC2, and D of Major Depressive Disorder; RC7 and Pt of Generalized Anxiety Disorder; RC8 and Sc of Psychotic Disorder; and RC9 and Ma of Bipolar

Disorder. Similarly, Sellbom et al. (2012) hypothesized the following RC Scales as predictors of SCID diagnosis: RCd and RC2 of Major Depressive Disorder, RC6 and RC8 of Schizophrenia, and RC9 of Bipolar Disorder (See SCID Diagnoses column of Table 2). Due to lack of a relevant SCID diagnostic category in the psychiatric sample, hypotheses were not made regarding RC3 and RC4, or their Clinical Scale versions (Hy and Pd, respectively), in prediction of any external criterion measure.

In prediction of PDF scale scores, scale selections were made following research by Barthlow et al. (1999) in which MMPI-2 Clinical Scales were selected as being conceptually related to specific PDF Scales. Barthlow and colleagues hypothesized the following Clinical Scale in prediction of PDF Scale: Hs of Somatic Symptoms; D of Depressed; and Pt of Anxious, Insecure, and Obsessive-Compulsive. The current study additionally selected the respective RC Scale version as predictor of each PDF Scale. Additional hypotheses were included using Clinical and RC Scale descriptions from Ben-Porath and Tellegen (2008), Graham (2006), and rational approach of the author. Pa and RC6 were hypothesized to predict Suspicious, and Ma and RC9 were hypothesized to predict both Narcissistic and Agitated (see PDF Scale column in Table 2).

Because SRF Scales were created to assess constructs similar to those of each MMPI-2 Clinical Scale, the Clinical Scale was selected as predictor of the respective SRF Scale. RC Scale counterparts were then chosen as predictors of individual SRF Scales as well. Only the SRF Scales matching Clinical Scales examined in the current study were selected (see SRF Scale column in Table 2).

Hypotheses

Previous research has shown most RC Scales do provide similar to greater convergent and discriminant validity than respective Clinical Scales despite having fewer items (Sellbom & Ben-Porath, 2005; Sellbom, Ben-Porath, et al., 2006; Simms et al., 2005). Enhanced utility of the RC compared to Clinical Scales, researchers explain, is a result of the process through which the

Table 2 *Scale Combinations Used in Testing Convergent, Discriminant, and Incremental Validity Hypotheses: Predicting SCID Diagnosis, PDF, or SRF Scale Scores using Selected Clinical and RC Scales as Predictors*

Predictors		Criteria		
		Psychiatric sample		Normative Couple sample
CS	RC	SCID Diagnostic Status[a]	PDF Scale(s)	SRF Scale
D	RCd	MDD	Depressed	D_R
Hs	RC1	Somatoform Disorder	Somatic Symptoms	Hs_R
D	RC2	MDD	Depressed	D_R
Pa	RC6	Schizophrenia	Suspicious	Pa_R
Pt	RC7	Anxiety Disorder	Anxious, Insecure, Obs-Comp	Pt_R
Sc	RC8	Schizophrenia	Psychotic Symptoms	Sc_R
Ma	RC9	Bipolar Disorder	Narcissistic, Agitated	Ma_R

Note. CS = Clinical Scale; RC = Restructured Clinical Scale; SCID = Structured Clinical Interview for DSM-IV; PDF = Patient Description Form; SRF = Spouse Rating Form; MDD = Major Depressive Disorder; RCd = Demoralization; RC1 = Somatic Complaints; RC2 = Low Positive Emotions; RC6 = Ideas of Persecution; RC7 = Dysfunctional Negative Emotions; RC8 = Aberrant Experiences; RC9 = Hypomanic Activation; Hs = Hypochondriasis; D = Depression; Pa = Paranoia; Pt = Psychasthenia; Sc = Schizophrenia; Ma = Hypomania; Obs-Comp = Obsessive-Compulsive; D_r = Depression; Hs_R= Hypochondriasis; Sc_R = Schizophrenia; Pt_R= Psychasthenia; Ma_R = Hypomania.
[a] Diagnostic status is a binary variable where a specific diagnosis being present was coded 1 and the rest of the diagnoses were coded 0.

RC Scales were created which (a) removed the Demoralization factor and (b) increased focus assessing one as opposed to several core constructs (Sellbom & Ben-Porath, 2005), thereby enabling clinicians to determine if scale elevations are due to symptoms of general emotional distress or a specific group of psychological symptoms (Sellbom, Ben-Porath, et al., 2006).

Examination of trends reported in previous research across multiple types of criterion measures identifies patterns in scale utility in which certain RC Scales (RC6, RC8, and RC9) tend to provide greater, while others (RCd, RC1, RC2, and RC7) provide similar to weaker utility than respective Clinical Scales. As mentioned by Nichols (2006) and Simms et al. (2005), Clinical Scales were specifically designed to assess the heterogeneous symptoms characteristic of clinical diagnoses enabling them to capture a more complex set of symptoms than do respective RC Scales. RC Scales, on the other hand, were constructed through factor analysis, which enhances their utility in assessment of relatively homogeneous symptom constructs compared to respective Clinical Scales. However, the complexity and resulting quantity of symptoms necessary to effectively represent the construct of each Clinical Scale (e.g., Depression) may vary in degree across scales giving impetus to differences in utility when constructs are transferred to the more homogeneous RC Scales. Symptom constructs of Clinical Scales characterized by relatively clear-cut symptom constructs may have been easier to reduce into fewer items leading to RC Scales that assessed symptoms at least as well, if not better, than respective Clinical Scales. The reverse may be true for revision of Clinical Scales comprising more multifaceted symptom content in which revision to the respective RC Scale created a scale not performing as well as the respective Clinical Scale.

For example, several studies have reported RC6 to provide greater assessment utility than Pa (Sellbom & Ben-Porath, 2005; Sellbom, Ben-Porath, et al., 2006; Sellbom, Graham, et al., 2006). Although Pa has not been specifically mentioned as assessing straightforward symptom constructs, authors of the RC Scales were able to successfully capture the essence of Pa in fewer items by focusing scale content mainly on symptoms of paranoia. Researchers (Sellbom, Ben-Porath, McNulty, Arbisi, & Graham, 2006), including critics of the RC Scales (Nichols, 2006),

describe that relegating items containing symptoms of paranoia previously found on other scales, as well as removing unrelated item clusters (e.g., naïveté), worked especially well in RC6 to create a scale that better represented the construct of paranoid ideation than its Clinical Scale version.

Like reasoning can be applied to the success of RC8 to provide greater assessment utility than Sc (Sellbom & Ben-Porath, 2005; Sellbom, Ben-Porath, et al., 2006; Sellbom, Graham, et al., 2006). Nichols (2006) writes the revision process from Sc to RC8 formed a scale containing item content well balanced to measure hallucinations, as well as unusual perceptions and experiences. The result was a scale that could identify positive symptoms of schizophrenia better than Sc or any other Clinical or Content Scale. Simms et al. (2005) further reports that while Pa and Sc showed some overlap in symptoms of paranoia and unusual perceptions, the revision process from respective Clinical Scale versions appears to have alleviated some of this concern in that traits centered on eccentric perceptions, the key symptom cluster representative of Sc, show stronger relations to RC8, while symptoms of paranoia are more closely related to RC6. Therefore, RC8 is an effective assessment of symptoms of Schizophrenia due to the increased focus on item content measuring symptoms related to unusual experiences and removal of content related to paranoia.

An increase in the proportion of item content assessing symptoms of aggression may explain a large part of the improvement in assessment utility of RC9 than Ma (Sellbom & Ben-Porath, 2005; Sellbom, Ben-Porath, et al., 2006; Sellbom, Graham, et al., 2006). Nichols (2006) reports that 71% of items contained in RC9 are items not found in the original Ma scale. Most of these new items comprise aggressive content, which appears to shift scale interpretation away from symptoms of mania not related to aggression. The finding that many cited studies have

found this increase in aggressive and hostile content to have improved the utility of RC9 above that of Ma implies this modification may not be as injurious as Nichols has proposed.

Conversely, research has shown weaker assessment utility for both RCd (Sellbom & Ben-Porath, 2005; Sellbom, Ben-Porath, et al., 2006) and RC2 (Sellbom, Ben-Porath, et al., 2006; Sellbom, Graham, et al., 2006; Simms et al., 2005) than for D. As compared to traits characteristic of Paranoia, Schizophrenia, and Mania, symptoms of Depression may be more difficult to assess with a narrower domain of item content due to the quantity of symptoms comprising this construct. Sellbom, Graham, et al. (2006) explain symptoms of Depression encompass elements of both negative and positive emotionality. RC2 measures only positive emotionality, while D assesses traits of both. Due to its assessment of only one symptom cluster, RC2 naturally would not perform as well as D in capturing symptoms of Depression. Similarly, Simms et al. (2005) demonstrated strong correlations of RCd with both positive and negative emotionality, but more so with negative emotionality. Though not explicitly stated by Sellbom, Graham, et al., it is possible to apply their logic to the inconsistent performance of RCd across studies: utility of RCd, as compared to D, was limited in ability to capture the full nature of symptoms of Depression due to its slight weighting towards symptoms of negative emotionality.

The complexity of symptoms captured in D differentially impacts the convergent and discriminant validity with respective RC Scales (RCd and RC2). Because D captures a more full range of symptoms, its average correlation with conceptually unrelated criterion measures (r_{mean}), as well as with conceptually relevant criterion measures (r) should be stronger than RCd or RC2. This increases the likelihood that symptoms assessed by D will covary with symptoms represented in multiple other Clinical Scales. In other words, even if the correlation of conceptually relevant criterion measure (r) is stronger with D than with RCd or RC2, the value

25

of r_{mean} will also be larger, making the difference between these two values (r_{dis}) similar to that of RCd or RC2. As a result, convergent, but not necessarily discriminant validity, of conceptually relevant criterion measures should naturally be greater for D than for RC2 or RCd.

The argument regarding construct complexity can also explain the weaker assessment utility of RC7 of symptoms of anxiousness or diagnosis of anxiety disorder than Pt (Sellbom & Ben-Porath, 2005; Sellbom, Ben-Porath, et al., 2006; Sellbom, Graham, et al., 2006; Simms et al., 2005). Simms et al. (2005) proposed a large proportion of item content of RC7 is focused around symptoms of negative emotionality, a characteristic of generalized anxiety disorder (GAD). Very few items of RC7 describe specific traits of anxiety disorders such as symptoms of anxiousness for panic disorder or reliving past symptoms for posttraumatic stress disorder (PTSD). By failing to include items representing symptoms characteristic of particular anxiety disorders, along with negative emotionality typical of GAD, authors of the RC Scales limited the ability of RC7 to gauge symptoms of anxiousness compared to Pt.

In the case of RC1, studies have generally shown RC1 and Hs demonstrated similar convergent and discriminant validity (Sellbom & Ben-Porath, 2005; Sellbom, Ben-Porath, et al., 2006; Simms et al., 2005), while RC1 yielded better predictive utility than Hs. The strong correlations of RC1 with symptoms of depression and anxiety, along with the expected somatic complaints shown by previous research (Sellbom, Ben-Porath, et al., 2006; Simms et al., 2005), may provide an explanation for this irregularity. Simms and colleagues explain the current models of anxiety and depressive disorders also include somatic symptoms (e.g., poor sleep), which results in natural overlap in symptoms among these constructs. The strong relations of RC1 with symptoms other than somatic complaints most likely limited convergent and discriminant validity while not impairing predictive utility.

26

This study formed hypotheses acknowledging differences in assessment utility across RC Scales as to their convergent validity (Hypotheses 1a-1g), discriminant validity (Hypotheses 2a-2g), and incremental utility (Hypotheses 3a-3g). This method, suggested by Archer (2006), provides more constructive information on scale performance to RC Scale users than research dictating widespread dismissal or approval for either group of RC or Clinical Scales. The following hypotheses are made regarding strength of zero-order correlations of conceptually relevant critiera with RC Scales and respective Clinical Scales:

H1a: RCd will show weaker correlations than D.

H1b: RC1 will show similar correlations as Hs.

H1c: RC2 will show weaker correlations than D.

H1d: RC6 will show stronger correlations than Pa.

H1e: RC7 will show weaker correlations than Pt.

H1f: RC8 will show stronger correlations than Sc.

H1g: RC9 will show stronger correlations than Ma.

The following hypotheses are made of discriminant validity of RC Scales of conceptually relevant criteria with discriminant validity (r_{dis}) calculated by subtracting the (a) average correlation with all conceptually unrelated criterion measures (r_{mean}) from the (b) correlation with conceptually relevant criterion measure (r). Larger (positive) differences between correlation coefficients indicated better discriminant validity.

H2a: RCd will show similar discriminant validity as D.

H2b: RC1 will show similar discriminant validity as Hs.

H2c: RC2 will show similar discriminant validity as D.

H2d: RC6 will show greater discriminant validity than Pa.

H2e: RC7 will show weaker discriminant validity than Pt.

H2f: RC8 will show greater discriminant validity than Sc.

H2g: RC9 will show greater discriminant validity than Ma.

The following hypotheses are proposed in prediction of conceptually relevant SCID diagnostic status, PDF, and SRF Scales:

H3a: Incremental utility of RCd to D will be less than that of D to RCd.

H3b: Incremental utility of RC1 to Hs will be greater than that of Hs to RC1.

H3c: Incremental utility of RC2 to D will be less than that of D to RC2.

H3d: Incremental utility of RC6 to Pa will be greater than that of Pa to RC6.

H3e: Incremental utility of RC7 to Pt will be less than that of Pt to RC7.

H3f: Incremental utility of RC8 to Sc will be greater than that of Sc to RC8.

H3g: Incremental utility of RC9 to Ma will be greater than that of Ma to RC9.

With the removal of the factor of demoralization from each Clinical Scale in the revision to the respective RC Scale, it is logical to assume that a portion of predictive utility was also taken away. To assess impact of this content removal on incremental utility for each individual RC Scale, exploratory analyses will be performed with both the hypothesized RC Scale and RCd entered in the same step as predictors of conceptually related criterion measures. Since no previous research was available using these techniques (to the author's knowledge), these analyses are exploratory. No specific hypotheses are made. The proportion of incremental validity the respective RC Scale, with the entry of RCd in the same step, will add above that of the Clinical Scale version will be compared to that the Clinical Scale adds above that of the entry of respective RC Scale and RCd in prediction of rationally selected SCID diagnostic status, PDF, and SRF Scales.

Table 3. *Study Hypotheses*

	Convergent Utility		Discriminant Utility		Incremental Utility
H1a	RCd will show weaker correlations with conceptually relevant criteria than D.	H2a	RCd will show similar discriminant validity of conceptually relevant criteria as D.	H3a	Incremental utility of RCd to D will be less than that of D to RCd in prediction of conceptually relevant criteria.
H1b	RC1 will show similar correlations….as Hs.	H2b	RC1 will show similar discriminant validity as Hs.	H3b	Incremental utility of RC1 to Hs will be greater than that of Hs to RC1…
H1c	RC2 will show weaker correlations…as D.	H2c	RC2 will show similar discriminant validity as D.	H3c	Incremental utility of RC2 to D will be less than that of D to RC2…
H1d	RC6 will show stronger correlations…as Pa.	H2d	RC6 will show greater discriminant validity than Pa.	H3d	Incremental utility of RC6 to Pa will be greater than that of Pa to RC6…
H1e	RC7 will show weaker correlations…as Pt.	H2e	RC7 will show weaker discriminant validity than Pt.	H3e	Incremental utility of RC7 to Pt will be less than that of Pt to RC7…
H1f	RC8 will show stronger correlations…as Sc.	H2f	RC8 will show greater discriminant validity than Sc.	H3f	Incremental utility of RC8 to Sc will be greater than that of Sc to RC8…
H1g	RC9 will show stronger correlations…as Ma.	H2g	RC9 will show greater discriminant validity than Ma.	H3g	Incremental utility of RC9 to Ma will be greater than that of Ma to RC9…

CHAPTER III

METHODOLOGY

Participants and Procedure

Sample I: Clinical Sample

Patients. Clinical sample consisted of inpatients and outpatients of the Samsung National Hospital in Seoul, Korea. Patients were administered tests in the following order: (1) Korean version of Structured Clinical Interview (SCID), (2) Korean versions of SCL-90-R (abbreviated form) and Korean MMPI-2, and (3) therapist ratings including a Medical Record Form, Mental Status, and Patient Description Form (PDF). The self-report measures (SCL-90-R and MMPI-2) were completed over a period of one to two days. Patients were diagnosed with SCID for DSM-IV Axis I Disorders (First, Spitzer, Gibbon, & Williams, 1997) with the exception of five individuals who were diagnosed with the Mini International Neuropsychiatric Interview (MINI).

Of the initial sample of 400 patients, 5 were removed due to exclusion criteria: Cannot Say $\geq$ 30, VRIN (T-score) $\geq$80, TTRIN (T-score) $\geq$80, and Fp >100. As shown in Table 4, the final clinical sample consisted of 176 men (M_{age} = 35.23, SD = 14.94) and 219 women (M_{age} = 40.72, SD =15.13). Nearly half of the patients reported never being married (n = 194 or 49.1%) while slightly less than half reported being married (n = 175 or 44.3%). Nearly one fourth of the patients reported being currently employed either full or part-time (n = 90 or 22.8%). Another fourth of the patients reported being students (n = 85 or 21.5%). Housewives made up another fourth of the sample (n = 100 or 25.3%). Most patients reported being graduates of either college (n = 73 or 18.5%) or high school (n = 112 or 28.4%). On average, patients had completed 13.82 years of education (SD = 3.32).

30

Patients were diagnosed with primary diagnosis of Schizophrenia ($n = 71$), Major Depressive Disorder (MDD) ($n = 85$), Bipolar Disorder ($n = 89$), Anxiety Disorder ($n = 86$), or Somatoform Disorder ($n = 64$). Roughly a quarter of these cases (26%) were diagnosed with a secondary diagnosis ($n = 102$) while less than 1% were diagnosed with a tertiary diagnosis ($n = 2$) (see Appendix A). Patients with primary diagnosis of Somatoform Disorder had the highest percentage (72%) of co-morbid diagnoses, while the overall most common co-morbid diagnoses were MDD with Anxiety Disorder ($n = 38$ or 10%). Only two patients with primary diagnosis of Schizophrenia had a co-morbid diagnosis. Patients with primary diagnosis of Anxiety Disorder or MDD also had a large percentage of co-morbid diagnoses (28% and 21%, respectively).

Table 4. *Descriptive Statistics of Sex and Age by Diagnostic Status*

			Age
Diagnostic Group		*f (%)*	*M (SD)*
Schizophrenia	Men	36 (9.1)	31.19 (10.34)
	Women	35 (8.8)	31.74 (11.77)
	Combined	71 (17.9)	31.46 (10.99)
Major Depressive Disorder	Men	32 (8.1)	38.31 (16.37)
	Women	53 (13.4)	44.08 (15.08)
	Combined	85 (21.5)	41.91 (15.74)
Bipolar Disorder	Men	34 (8.6)	28.06 (10.59)
	Women	55 (13.9)	33.96 (12.35)
	Combined	89 (22.5)	31.71 (11.99)
Anxiety Disorder	Men	51 (12.9)	40.69 (16.50)
	Women	35 (8.9)	47.54 (15.23)
	Combined	86 (21.8)	43.48 (16.26)
Somatoform Disorder	Men	23 (5.8)	36.13 (16.31)
	Women	41 (10.4)	47.24 (14.07)
	Combined	64 (16.2)	43.25 (15.73)
Total	Men	176 (44.6)	35.28 (14.99)
	Women	219 (55.4)	40.71 (15.13)
	Combined	395	38.29 (15.29)

Therapists. Fifteen therapists (13 women and 2 men) completed the SCID, therapist rating form, medical record form, mental status examination. All therapists were blind to the

purpose of the study and MMPI-2 results. One therapist completed all forms for each case with each therapist rating multiple cases (number of cases: range = 3 to 71, *mdn* = 21). All therapists held Master's degree and were receiving training as Clinical Psychologists. Due to no two clinicians rating the same profile, no interrater reliability data were available. Clinicians were under supervision for his or her supervisor; however, diagnosis used was the original diagnosis from the clinician prior to receiving feedback for the supervisor. As clinician training procedures were in use, it was assumed that adequate levels of diagnostic reliability were achieved.

Sample II: Korean Normative Couple Sample

The second sample consisted of a portion of the normative sample from the Korean MMPI-2 standardization project (Han, 1993). Potential participants were contacted through telephone and were invited to the study. Testing sessions were conducted in designated testing centers in groups of approximately 25 -50 participants. Individuals who participated with their spouse as a married couple completed the Korean MMPI-2, a Biographical Information Form, and a Life Events Form. Each couple was then instructed to rate his or her spouse on the Spouse Rating Form and received 10,000 won (roughly 10 U.S. dollars) for participation.

The initial normative sample consisted of 1,587 individuals, of which 235 were removed based on the following criteria: incomplete responses on Biographical Information or Life Event Forms, double answers on MMPI-2 items ($n_{removed}$ = 59). Further cases ($n_{removed}$ = 154) were removed for failure to meet the following: Cannot say > 0, VRIN > 17, TRIN < 4, TRIN > 15, or KI (Korean Infrequency Scale) > 21. Further participants were randomly removed to ensure appropriate representativeness ($n_{removed}$ = 22). Of these remaining cases, 256 participants completed the study as a married couple (128 couples). From this sample, 25 participants were removed based on the before mentioned criteria.

The final sample consisted of 231 married adults, 216 of which completed the study as a married couple (108 couples). Men averaged 43.06 years of age ($SD = 8.80$) and 12.30 years of education ($SD = 3.15$), while women were slightly younger ($M_{age} = 40.43$, $SD = 8.75$) and had less education ($M = 11.42$, $SD = 2.67$). Couples reported being married an average of 15.02 years ($SD = 9.83$). Nearly one third of the participants reported being currently employed ($n = 62$ or 28.7%) or employed part-time ($n = 14$ or 6.5%). The majority of participants reported no religious affiliation ($n = 95$ or 44.0%). Of the remaining participants, the most frequently endorsed religions were Christianity ($n = 58$ or 26.9%) and Buddhism ($n = 37$ or 17.1%). The most frequently endorsed occupation was response option "none of the above" ($n = 111$ or 51.4%), followed by Clerical worker ($n = 42$ or 19.4%), and Skilled craftsperson ($n = 22$ or 10.2%). Most individuals reported income range between 20,000,000 – 29,999,999 won ($18,182 - $27,272; $n = 71$ or 32.9%), followed by 30,000,000 to 39,999,999 won ($27,273 - $36,363; $n = 46$ or 21.3%), and 10,000,000 to 19,999,999 won ($36,364 – 18,181; $n = 41$ or 19.0%).

<center>Measures</center>

Korean MMPI-2/ MMPI-2-RF

The MMPI-2 (Butcher et al., 1989) was initially translated into Korean by Han (1993, 1996). The preliminary version of the Korean MMPI-2 was created via multi-stages of translation and back-translation (see Han, 1993, 1996). This preliminary MMPI-2 was piloted in a sample of 726 Korean college students from eight universities (1993, 1996). Items that produced unanticipated endorsement frequencies were retranslated. This revised version was evaluated by a professional translation company that was requested by the University of

<center>33</center>

Minnesota Press to check the accuracy of the translation, and 50 items determined to have minor

problems were retranslated. In 2002, slight modifications to 179 items were made by the Korean

MMPI-2 Standardization Committee, as it had been approximately 10 years since the initial

translation. Thirteen of these items were deemed significantly altered from the original

translation; these items were submitted to a back translation into English. After additional

translation and back-translation, the final version of the Korean MMPI-2 was determined (Kim et

al., 2005). Han and her colleagues (2005) conducted investigation of the psychometric

properties of the Korean MMPI-2 and determined that they were comparable to those obtained

with the American MMPI-2. For a detailed description of adaptation procedures, see Han (1993,

1996) and Kim et al. (2005).

The MMPI-2-RF is a "broad-spectrum" self-report inventory consisting of 338 true or

false statements on personality and psychopathology (Tellegen & Ben-Porath, 2008, p. 1). All

MMPI-2-RF items can be derived from the MMPI-2. A total of 338 items make up the 50 scales,

which are divided into 8 validity scales, 9 Restructured Clinical (RC) Scales (Tellegen et al.,

2003), and 33 scales measuring a wide range of psychological maladjustment, specific problem,

interest, and personality traits, 23 of which include the Specific Problem Scales. For both the

U.S. and Korea, the normative sample of the MMPI-2 was also used for MMPI-2-RF with minor

modifications. The Korean MMPI-2-RF was published in 2011 (Han et al, 2011). For this

study, the 567-item Korean MMPI-2 (Kim et al., 2005) was administered to participants. The

MMPI-2- RF scales were scored based on the MMPI-2 items. RC Scales are scored in both the

MMPI-2 and MMPI-2-RF. All Korean MMPI-2-RF normative samples (651 men and 651

women) are in the MMPI-2 normative sample (701 men and 651 women).

For the clinical sample, internal consistency (Cronbach's alpha) coefficients for RC and Clinical Scales were good ranging between .81 (RC2) to .89 (RCd) (*mdn* = .84) and between .70 (Ma) to .92 (Sc) (*mdn* = .81), respectively. Coefficients of the current study were similar to those listed for the RC Scales in the U.S. MMPI-2-RF manual for both outpatient and inpatient samples. For the normative couple sample, values were slightly lower with acceptable to good values ranging between .69 (RC2) and .83 (RCd) (*mdn* =.75) for the RC Scales and .57 (Pa) to .88 (Sc) (*mdn* = .75) for the Clinical Scales. These values were similar to those listed for the MMPI-2-RF normative sample in the manual.

Structured Clinical Interview for Diagnosis on DSM- IV (SCID)

Inpatients and outpatients received psychiatric diagnosis based on the fourth edition of the Diagnostic and Statistical Manual of the Mental Disorders (DSM-IV; First et al., 1997). The SCID assesses current and lifetime psychiatric disorders in a structured interview format using DSM-IV diagnostic criteria. This interview is designed to make Axis I diagnoses. The results of the interview present a record of the presence or absence of symptoms for each of the disorders for current or lifetime occurrence. Using specific criteria for the number of symptoms present, the clinician is able to categorize clients into one or more than one Axis I disorder by means of asking structured questions along the way of the interview.

Patients were diagnosed based on Axis I disorders including Schizophrenia, Major Depressive Disorder (MDD), Bipolar Disorder, Anxiety Disorder, Somatoform, or none of the above. Patients could be classified with multiple diagnoses (e.g., MDD and Anxiety Disorder), but each group was independent of the other groups. The primary diagnosis was used to categorize patients (see Appendix A for frequencies and details of pure vs. co-morbid diagnoses). Despite the present of co-morbid diagnoses, and in some diagnoses to a large

proportion (e.g., current sample Somatoform Disorder has 72% co-morbidity), past research has indicated co-morbid diagnoses do not decrease the validity of the study. As long as the goals are to differentiate patients with one diagnosis (e.g., Schizophrenia) from all other patient diagnostic groups, co-morbid data may be used (e.g., Wetzler et al., 1998).

Patient Description Form (PDF)-Therapist Rating Form

The Patient Description Form (PDF; Graham, Ben-Porath, & McNulty, 1999) contains 188 items and 25 scales. Therapists rated items using a 5-point scale (not at all, slight, moderate, high, and very high) to describe personality traits and symptoms of patients. The 25 scales were created using both a rational and statistical approach. Han initially translated these 188 items into Korean and two bilingual psychologists (Jeeyoung Lim and Sangil Kwon) reviewed Han's translation. After several revision translations, a Korean version of the PDF was constructed. The length was shortened due to the Korean MMPI-2 Standardization Project Committee's concern over the item length of the PDF. An abbreviated version was created to form the Therapist Rating Form. Han reviewed the items in each scale (a total of 25 scales) and selected 2-5 items judged to be the most salient behaviors/symptoms of each scale domain. The final Therapist Rating Form contained 70 items and 23 scales. Han selected 70 items that were most representative of the behavior or symptoms for each scale.

Nine of the 25 scales were selected for use as criterion measures in the current study. These scales (with number of items) included Anxious (3), Insecure (3), Obsessive-Compulsive (4), Depressed (5), Somatic Symptoms (4), Psychotic Symptoms (5), Suspicious (3), Narcissistic (4), and Agitated (3). The internal consistency (Cronbach's alpha) coefficients in this study ranged from .53 (Anxious) to .90 (Psychotic Symptoms) (*mdn* = .81).

36

Spouse Rating Form (SRF)

The Spouse Rating Form consisted of 70 items, 60 of which were designed to measure symptoms similar to that of the 10 MMPI-2 Clinical Scales. Each scale comprised six items. The 60 items were selected from the Spouse Rating Form which was used as part of the American MMPI Restandardization Project (Butcher et al., 1989) and from the MMPI-2 Clinical Scale descriptors listed in Graham (2000). An additional 10 items were created by Han as potential behavioral correlates of the clinical scales. These 70 items were translated into Korean by Han and Lim, and the translation of these items was checked for accuracy by the Korean MMPI-2 Standardization Committee. The following were selected for the current study: Hypochondriasis (Hs_R), Depression (D_R), Paranoia (Pa_R), Psychasthenia (Pt_R), Schizophrenia (ScR_), and Hypomania (Ma_R). The internal consistency (Cronbach's alpha) coefficients in this study ranged from .30 (Ma_R) to .70 (Pa_R) (*mdn* = .46). Although Graham (2006) listed several items as descriptors of multiple Clinical Scales, current study authors selected items as markers of the scale that were most prototypical. For example, item 59 (Has a wide range of interests) was categorized as marker of Ma, despite its being listed as describing elevated scores on Psychopathic Deviate (Pd) and Sc.

Analyses

Data Screening

Missing values. Missing values for each item of the Clinical and RC scales were examined. Details of this process are presented in Appendix B to clarify reading of paper.

Equivalence of groups. In the clinical sample, diagnostic status groups were tested for equivalence in terms of demographic variables using one-way ANOVAs (i.e., age) and Chi-

square analyses (i.e., gender, major, religion). In the Korean normative couple sample, men and women were tested for differences in demographic variables (e.g., age, education, and religion). For all hypothesis tests, raw scale scores were used.

Outliers: Logistic regression. Analysis of the logistic regression model standardized residuals was conducted predicting each of the five diagnostic status conditions versus all other diagnoses. This included a total of 14 hierarchical logistic regression analyses using the hypothesized Clinical and RC Scale as predictors of each of the five diagnoses (e.g., two predicting Somatoform Disorder using Hs/RC1, two predicting Major Depressive Disorder using D/RCd, two predicting Major Depressive Disorder using D/RC2, two predicting Schizophrenia from Pa/RC6, two predicting Schizophrenia from Sc/RC8, two predicting Anxiety Disorder from Pt/ RC7, and two predicting Bipolar Disorder from Ma/RC9). This also included another 14 hierarchical logistic regression analyses using the hypothesized Clinical and RC Scale with the addition of RCd as predictors of each of the five diagnoses. Standardized residuals with z-scores $\geq \pm 3.29$ were marked as potential outliers and checked for data entry error (Tabachnick & Fidell, 2007).

Outliers: Multiple regression. Outliers using regression model standardized residuals and Cook's Distance values were assessed. Cases with a standardized residual $\geq \pm 3.29$ (Tabachnick & Fidell, 2007) or Cook's Distance values > 1.0 (Stevens, 1984) or appearing disconnected from the rest of the data were marked and checked for data entry errors.

Multicollinearity: Logistic regression. Performing hierarchical logistic regression requires absence of multi-collinearity (Tabachnick & Fidell, 2007). Logistic regression model parameters and standard errors were examined for extremely large values, which indicate highly correlated predictor variables (multicollinearity).

Multicollinearity and singularity: Multiple regression. Multiple regression analysis also assumes absence of multicollinearity and singularity. Indications of multicollinearity among the independent predictor variables were assessed using the collinearity diagnostics function in SPSS Version 17. Condition index values greater than 30 and variance proportions greater than .50 for two different variables or more indicate a serious problem in collinearity (Tabachnick & Fidell, 2007). Tolerance was used to assess singularity. Tolerance (1 - SMC where SMC is squared multiple correlation) in which values close to zero indicate singularity.

Linearity: Logistic regression. Logistic regression models assume a linear relationship between the predictor variable and the logit transformation of the dependent variable. This assumption was tested via the Box-Tidwell method. Each predictor variable was entered in the first step and an interaction term between each predictor and its natural logarithm was entered in the second step of the regression model (Tabachnick & Fidell, 2007). A significant interaction term indicates the assumption is violated and a variable transformation may be considered (e.g., logrhythmic). In prediction of each diagnostic status, logistic regression analyses were performed with the hypothesized Clinical and RC Scale entered in step one, and the two interaction terms (Clinical Scale x its natural log and RC Scale x its natural log) were entered in step two. For example, D and RCd were entered in step one, and D x *ln* (D) and RCd x *ln* (RCd) were entered in step two.

Normality, linearity, and homoscedasticity: Multiple regression. Presence of normality, linearity, and homoscedasticity of residuals is assumed in multiple regression models. These assumptions were assessed by examining residuals of scatterplots between the criterion (dependent variable) and the errors of prediction. Models that follow these assumptions should appear normally distributed around dependent variable scores, residuals should show a straight-

line relationship with the predicted scores, and the variance of the residuals around the dependent variable scores should be relatively similar across all predicted score (Tabachnick & Fidell, 2007).

Independence of error. Due to the fact that external criterion measures (SCID and PDF) for the clinical sample were completed by a relatively limited number of therapists (15), concern was present for the impact of differences between therapists on client ratings. Past research has shown therapist variables (e.g., type or length of training) can have substantial impact on treatment of their clients accounting for between 5% (e.g., Kim, Wampold, & Bolt, 2006; Lutz, Leon, Martinovich, Lyons, & Stiles, 2007; Wampold & Brown, 2005) and 30% (Larrison & Schoppelrey, 2011) of the variability in client outcome variables (e.g., ratings of perceived levels of distress or functioning). Clients are therefore more likely to show greater level of similarities with clients from the same therapist than with clients from a different therapist (Adelson & Owen, 2012). This degree of relatedness among clients of the same therapist violates the assumption of independence of errors. Researchers (e.g., Adelson & Owen, 2012; Minami, Brown, McCulloch, & Bolstrom, 2012; Okiishi, Lambert, Nielsen, & Ogles, 2003; Wampold & Brown, 2005) have proposed use of multilevel modeling (MLM) in cases such as these, because it does not require independence of errors as do hierarchical and hierarchical logistic regression. In MLM, data are arranged hierarchically (e.g., clients are nested within therapists) allowing for intercepts (means) and slopes (relationships between the independent and dependent variables) to vary at higher levels. In the case of the current study, the relationship between therapist ratings and RC or Clinical Scales is allowed to vary between therapists.

Due to the limited number of therapists included in the current dataset (15 with several providing ratings of as few as three clients), results from MLM will be interpreted cautiously.

40

When MLM is performed with as few as 10 therapists, tests for differences between therapists may be biased in that results are more likely to show significant differences between therapists even when there is no difference (Type I error). Including 30 or more therapists is recommended (Adelson & Owen, 2012). In addition, the cutoff number of clients per therapist has ranged from four clients (Larrison & Schoppelrey, 2011; Wampold & Brown, 2005) to 30 clients (Okiishi, Lambert, Eggett, Nielsen, & Dayton, 2006) who suggest 30 as a cutoff, saying 14 was too few for generalizability. MLM will therefore be used as a supplementary tool to verify results of hierarchical logistic and hierarchical regression analyses do not differ substantially when accounting for degree of relationship between clients of the same therapists.

Testing Hypotheses

Convergent validity. To examine convergent validity (Hypotheses H1a- H1g), zero-order correlations were calculated between each of Clinical and RC Scales and conceptually relevant criterion scale from SCID diagnoses, PDF Scales, and SRF Scales. The difference in strength of correlation of each Clinical and respective RC Scale version with conceptually relevant criterion measure was then assessed using two criteria. Although no correlation magnitude has been explicitly identified in previous research, present author examined patterns of language used in interpreting meaningful differences in convergent validity of RC with that of Clinical Scales from previous studies. This examination indicated an estimated correlation difference of $> .05$ (Sellbom & Ben-Porath, 2005; Simms et al., 2005). Secondly, previous research (Sellbom, Ben-Porath, & Graham, 2006) has also tested the difference between dependent correlations (correlations that share a common variable) using a z test (Steiger, 1980). Therefore, in order for zero-order correlations of RC or Clinical Scale with conceptually relevant criterion measures to signify greater convergent validity, the following criteria must be met: (a) differences in

correlation strength must be > .05 and (b) difference in correlation strength must be significant indicated by a *z* test (*p* < .05). For example, to test convergent validity of RCd and D, both scales were correlated with each of the three external criterion measures: (a) binary diagnostic status (Major Depressive Disorder vs. other diagnosis; SCID), (b) Depressed (PDF), and (c) Depression (SRF). The difference between the two zero order correlations was then evaluated.

Discriminant validity. Discriminant validity (Hypotheses H2a - H2g) coefficients were obtained by subtracting the average zero-order correlation with all conceptually unrelated criterion measures (r_{mean}) from the zero-order correlation with conceptually relevant criterion measure (*r*). Larger (positive) differences between correlation coefficients indicated meaningfully greater discriminant validity based on same criterion described for convergent validity analyses (> .05). For example, to test discriminant validity of RCd and D, both scales were correlated with each of the three external criterion measures: (a) binary diagnostic status (Major Depressive Disorder vs. other diagnosis; SCID), (b) Depressed scale (PDF), and (c) Depression (SRF). Then, RCd and D were correlated with all conceptually unrelated criterion measures: (a) binary diagnostic status (Somatoform Disorder vs. other diagnosis, Schizophrenia vs. other diagnosis, Anxiety Disorder vs. other diagnosis, and Bipolar Disorder vs. other diagnosis; SCID), (b) Somatic Symptoms, Suspicious, Anxious, Insecure, Obsessive-compulsive, Psychotic Symptoms, Narcissistic, and Agitated (PDF), and (c) (Hs_R, Pa_R, Pt_R, Sc_R, and Ma_R (SRF). An average was then calculated. The difference between the average correlation and the zero order correlation was then evaluated separately for each criterion measure.

Incremental utility. To examine the incremental utility of each RC scale to its corresponding Clinical Scale (Hypotheses H3a – H3g), regression analyses were performed in which the Clinical Scale was entered in step one, and the RC was entered in step two. Regression

analyses were also conducted with the order of entry reversed: the RC Scale was entered in step one, and the Clinical Scale was entered in step two. For a regression model with a binary criterion (diagnostic status from SCID), a hierarchical logistic regression analysis was used, whereas a hierarchical multiple regression was used for a model with a continuous variable (PDF and SRF). For example, two hierarchical logistic regression analyses were performed to examine the incremental predictive validity of RC1 and Hs in prediction of Somatoform Disorder (Hypothesis H3b). In the first regression analysis, RC1 was entered in the first step, and Hs was entered in the second step. In the second regression analysis, Hs was entered in the first step, and RC1 was entered in the second step.

Hypotheses of incremental utility predicting PDF Scales were tested by performing three sets of hierarchical regression analyses to examine incremental predictive validity of Clinical and respective RC Scales in prediction of conceptually related PDF scales (see Table 2 for explanation of related scales). Both RC7/Pt and RC9/Ma were hypothesized to predict more than one PDF Scale. Multiple regression analyses were necessary using each separate PDF scale in a separate set of regression analyses. For example three sets of hierarchical regression analyses were conducted with Pt and RC7 predicting PDF Scales (1) Anxious, (2) Insecure, and (3) Obsessive-Compulsive. Similarly two sets of regression analyses were conducted predicting PDF scales Agitated and Narcissistic from Ma and RC9. Similar to hierarchical logistic regression models, analyses were conducted entering RCd in the same step with the hypothesized RC scale.

CHAPTER IV

RESULTS

Data Screening and Assumptions Checks

Before hypotheses were tested, specified assumptions were examined. These analyses

are presented in Appendix B to ease reading and clarity of paper.

Testing Hypotheses

Convergent Validity: Clinical Sample: Structured Clinical Interview for Diagnosis on DSM-IV (SCID)

As shown in Table 5, RC Scales demonstrated relatively similar to greater convergent

validity of a majority of SCID diagnoses than respective Clinical Scales. Schizophrenia showed

stronger correlations with RC8 ($r = .11$) than with Sc ($r = -.01$, Steiger's $z = 3.73$, $p < .001$), as

did diagnosis of Bipolar Disorder with RC9 ($r = .36$) than with Ma ($r = .29$, Steiger's $z = 2.26$, $p

< .05$), while diagnosis of Schizophrenia showed slightly stronger correlations with RC6 ($r = .16$)

than with Pa ($r = .12$, Steiger's $z = 1.16$, $p > .05$). Diagnosis of Somatoform Disorder yielded

relatively similar strength correlation with RC1 ($r = .17$) as with Hs ($r = .19$, Steiger's $z = -1.39$,

$p > .05$), as did Anxiety Disorder with RC7 ($r = -.03$) as with Pt ($r = -.04$, Steiger's $z = .42$, $p >

.05$). In contrast, diagnosis of MDD showed stronger correlation with D ($r = .22$) than with RCd

($r = .15$), though this difference was not significant (Steiger's $z = 1.49$, $p > .05$). Diagnosis of

MDD also showed slightly stronger correlation with D than with RC2 ($r = .17$; Steiger's $z =

1.45$, $p > .05$). Hypotheses H1b (RC1), H1f (RC8), and H1g (RC9) were supported. Although

results did follow expected patterns in comparative convergent validity, Hypotheses H1a (RCd),

H1c (RC2), and H1d (RC6) were partially supported, due to failure to meet either (a) magnitude

44

of correlation difference greater than .05, or (b) significance of difference in correlation strength indicated by Steiger's *z* test. Hypothesis H1e (RC7) was not supported.

Table 5. *Correlations of Structured Clinical Interview for Diagnosis (SCID) on DSM-IV Diagnostic Status with Clinical and Restructured Clinical (RC) Scales*

	MDD		SOM		SCZ		ANX		BP	
	CS	RC	CS	RC	CS	RC	CS	RC	CS	RC
RCd		.15^{H1a}		-.08		-.09		-.13		.13
1	.08	.02	.19^{H1b}	.17^{H1b}	-.22	-.18	-.06	-.09	.02	.08
2	.22^{H1a,c}	.17^{H1c}	.13	.08	-.15	-.05	.01	.01	-.21	-.21
3	.11	-.03	.20	-.07	-.15	-.13	-.01	.01	-.14	.21
4	.09	-.02	-.10	-.11	-.01	-.07	-.18	-.12	.19	.30
6	-.00	-.04	-.10	-.13	.12^{H1d}	.16^{H1d}	-.18	-.13	.15	.13
7	.11	.04	-.07	-.13	-.13	-.10	-.04^{H1e}	-.03^{H1e}	.11	.20
8	.05	-.06	-.12	-.13	-.01^{H1f}	.11^{H1f}	-.15	-.15	.21	.21
9	-.02	-.07	-.13	-.17	-.02	-.06	-.14	-.08	.29^{H1g}	.36^{H1g}

Note. Diagnostic status was coded: 1 = diagnosis present; 0 = other. CS = Clinical Scale; RC = Restructured Clinical Scale; MDD = Major Depressive Disorder; SOM = Somatoform Disorder; SCZ = Schizophrenia; ANX = Anxiety Disorder; BP = Bipolar Disorder. CS = Clinical Scale; RC = Restructured Clinical Scale; RCd = Demoralization; 1 = Hypochondriasis; RC1 = Somatic Complaints; 2 = Depression; RC2 = Low Positive Emotions; 3 = Hysteria; RC3 = Cynicism; r = Psychopathic Deviate; RC4 = Antisocial Behavior; 6 = Paranoia; RC6 = Ideas of Persecution; 7 = Psychasthenia; RC7 = Dysfunctional Negative Emotions; 8 = Schizophrenia; RC8 = Aberrant Experiences; 9 = Hypomania; RC9 = Hypomanic Activation.
[Hxx] Correlations used to evaluate hypotheses are indicated by their respective number and letter.

Discriminant Validity: Clinical Sample: Structured Clinical Interview for Diagnosis on DSM-IV (SCID)

Comparisons of discriminant validities (r_{dis}) of the Clinical and RC Scales of SCID diagnosis can be calculated from Table 5 by subtracting the (a) average correlation with all conceptually unrelated criterion measures (r_{mean}) from the (b) correlation with conceptually relevant criterion measure (*r*). For example, the correlation of RC8 with the conceptually relevant diagnosis of Schizophrenia is *r* = .11, while the average (absolute value) correlation of RC8 with all conceptually related diagnoses is .14, r_{mean} = ([MDD *r* = .06] + [Somatoform Disorder *r* = .13] + [Anxiety Disorder *r* = .15] + [Bipolar Disorder *r* = .21]) / 4 = .14. Therefore, discriminant validity (r_{dis}) of RC8 is -.03 ([*r* = .11] − [r_{mean} =.14] = -.03) (see Appendix C for complete calculations). RC8 did show greater discriminant validity of diagnosis of Schizophrenia

45

than did Sc (r_{dis} = -.14), while most other RC and respective Clinical Scales showed relatively similar discriminant validity of SCID diagnoses. For example, RC6 (r_{dis} = .05) showed slightly greater discriminant validity of diagnosis of Schizophrenia than Pa (r_{dis} = .01), as did RC9 of diagnosis of Bipolar Disorder (r_{dis} = .26) than Ma (r_{dis} = .21). Discriminant validities for both Clinical and RC Scale versions were nearly identical for RC1 (r_{dis} = .08) and Hs (r_{dis} = .09) of Somatoform Disorder, RC2 (r_{dis} = .08) and D (r_{dis} = .09) of diagnosis of MDD, and RC7 (r_{dis} = -.15) and Pt (r_{dis} = -.15) of diagnosis of Anxiety Disorder. Similar to comparisons of convergent validity, D (r_{dis} = .09) did show slightly greater discriminant validity of diagnosis of MDD than RCd (r_{dis} = .04). Hypotheses H2a (RCd), H2b (RC1), H2c (RC2), and H2f (RC8) were supported, and Hypotheses H2d (RC6) and H2g (RC9) were partially supported as they did follow predicted trends, but failed to meet magnitude of difference greater than .05. Hypothesis H2e (RC7) was not supported.

Convergent Validity: Clinical Sample: Patient Description Form (PDF)

Zero-order correlations of conceptually relevant Patient Description Form (PDF) Scales with RC and Clinical Scales are presented in Table 6. Comparing convergent validity of RC with respective Clinical Scales indicated mixed results: RC Scales provided similar to greater convergent validity of just over half of conceptually relevant PDF Scales than did Clinical Scale versions. Psychotic was more strongly correlated with RC8 than with Sc (rs = .28 and .09, respectively; Steiger's z = 5.69, $p < .001$). Depressed showed slightly stronger correlation with RCd (r = .56) than with D (r = .52; z = 1.06, $p > .05$), as did Agitated with RC9 (r = .21) than with Ma (r = .17; z = 1.30, $p > .05$). Both RC and Clinical Scales yielded similar strength correlations with Narcissistic (RC9 [r = .19] and Ma [r = .19], z = .09, $p > .05$) and Suspicious

Table 6. *Correlations of Patient Description Form (PDF) Scales with Restructured Clinical (RC) and Clinical Scales*

	Dep		Som		Susp		Anx		Insec		Obs-Co		Psych		Narc		Agi	
	CS	RC	CS	RC	CS	RC	CS	RC	CS	RC	CS	RC	CS	RC	CS	RC	CS	RC
RCd		.56[H2a]		.12		.09		.22		.41		.15		-.06		-.06		.04
1	.44	.38	.48[H2b]	.39[H2b]	-.05	.01	.20	.15	.15	.14	.00	-.04	-.19	-.11	-.01	.02	.09	.10
2	.52[H2a,c]	.46[H2c]	.39	.22	-.15	-.10	.31	.25	.29	.35	.20	.19	-.25	-.16	-.17	-.25	-.05	-.13
3	.33	.17	.46	.01	-.07	.08	.19	.09	.08	.08	.08	-.09	-.14	-.09	-.00	.07	.05	.13
4	.43	.24	.06	-.03	.18	.16	.12	.07	.34	.21	.03	-.07	.02	.02	.08	.15	.18	.17
6	.39	.20	-.00	-.09	.33[H2d]	.34[H2d]	.09	.04	.30	.19	.09	-.01	.27	.34	.06	.06	.19	.19
7	.51	.38	.17	.03	.07	.11	.26[H2e]	.20[H2e]	.35[H2e]	.28[H2e]	.12[H2e]	.08[H2e]	-.09	-.02	-.07	-.02	.07	.11
8	.49	.26	.07	-.02	.22	.29	.15	.04	.36	.15	.06	-.01	.09[H2f]	.28[H2f]	.04	.11	.14	.19
9	.08	.04	-.11	-.13	.19	.19	-.08	-.04	.02	.05	-.11	-.10	.14	.06	.19[H2g]	.19[H2g]	.17[H2g]	.21[H2g]

Note. Dep = Depressed; Som = Somatic Symptoms; Susp = Suspicious; Psych = Psychotic Symptoms; Anx = Anxious; Insec = Insecure; Obs-Co = Obsessive-Compulsive; Narc = Narcissistic; Agi = Agitated; CS = Clinical Scale; RC = Restructured Clinical Scale; RCd = Demoralization; 1 = Hypochondriasis; RC1 = Somatic Complaints; 2 = Depression; RC2 = Low Positive Emotions; 3 = Hysteria; RC3 = Cynicism; 4 = Psychopathic Deviate; RC4 = Antisocial Behavior; 6 = Paranoia; RC6 = Ideas of Persecution; 7 = Psychasthenia; RC7 = Dysfunctional Negative Emotions; 8 = Schizophrenia; RC8 = Aberrant Experiences; 9 = Hypomania; RC9 = Hypomanic Activation.
[Hxx] Correlations used to evaluate hypotheses are indicated by their respective number and letter.

(RC6 [$r = .34$] and Pa [$r = .33$]; $z = .35$, $p > .05$). In contrast, Somatic Complaints showed stronger correlations with Hs than with RC1 ($rs = .48$ and .39, respectively; $z = 5.66$, $p < .001$). Both Anxious ($rs = .26$ and .20, respectively; $z = 2.25$, $p < .05$) and Insecure ($rs = .35$ and .28, respectively; $z = 2.93$, $p < .01$) showed stronger correlations with Pt than with RC7, while Obsessive-Compulsive ($rs = .12$ and .08, respectively; $z = 1.66$, $p > .05$) showed slightly stronger correlation with Pt than with RC7. Depressed also showed stronger correlation with D ($r = .52$) than with RC2 ($r = .46$; $z = 1.89$, $p = .06$), though this difference was not significant. Hypothesis H1f (RC8) was supported, and H1e (RC7) was supported in two of three criterion measures with partial support of the third measure. Hypotheses H1c (RC2) and H1g (RC9) of Agitated followed predicted trends, but were partially supported due to failure to meet either (a) magnitude of correlation difference greater than .05, or (b) significance of difference in correlation strength indicated by Steiger's z test. Hypotheses H1a (RCd), H1b (RC1), H1d (RC6), and H1g (RC9) of Narcissistic were not supported.

Discriminant Validity: Clinical Sample: Patient Description Form (PDF)

As calculated from Table 6, RC Scales showed greater discriminant validity than Clinical Scale versions of several of the PDF Scales including: RCd of Depressed ($r_{dis} = .42$) than D ($r_{dis} = .29$), RC8 of Psychotic ($r_{dis} = .15$) than Sc ($r_{dis} = -.10$), and RC9 of Agitated ($r_{dis} = .11$) than Ma ($r = .05$) (see Appendix D for complete calculations). Most other comparisons indicated relatively similar discriminant validity among RC and Clinical Scales. For example, RC6 showed slightly greater discriminant validity of Suspicious ($r_{dis} = .20$) than did Pa ($r_{dis} = .16$). Both RC7 and Pt showed relatively equivalent discriminant validity of Anxious ($rs_{dis} = .07$ and .08, respectively), Insecure ($rs_{dis} = .16$ and .18, respectively), and Obsessive-Compulsive ($rs_{dis} = -.06$ and -.08, respectively), as did RC9 ($r_{dis} = .09$) and Ma ($r_{dis} = .08$) of Narcissistic. RC2 ($r_{dis} = .25$)

48

also demonstrated only slightly weaker discriminant validity of Depressed than did D ($r = .29$). The one exception was in discriminant validity of Somatic Symptoms in which RC1 ($r_{dis} = .27$) provided weaker discriminant validity of Somatic Symptoms than Hs ($r_{dis} = .34$).   Hypotheses H2c (RC2) and H2f (RC8) were supported, and H2g (RC9) was supported in only one of two conceptually relevant criterion measures. Hypothesis H2d (RC6) was partially supported as scales did follow predicted trends, but failed to meet magnitude of correlation difference greater than .05. Hypotheses H2a (RCd), H2b (RC1), and H2e (RC7) were not supported.

Convergent Validity: Korean Normative Couple Sample: Spouse Rating Form (SRF)

Table 7 shows the results of zero-order correlations of Spouse Rating Form (SRF) Scales with Clinical and RC Scales using the Korean normative couple sample. Similar to results of PDF Scales, RC Scales showed similar to greater convergent validity of nearly half of SRF Scales as did Clinical Scales. D_R was more strongly correlated with RCd ($r = .37$) than with D ($r = .19$, Steiger's $z = 2.76$, $p < .01$). Sc_R also showed stronger correlation with RC8

Table 7. *Correlations of Spouse Rating Form (SRF) Scales with Clinical and Restructured Clinical (RC) Scales*

	D_R		Hs_R		Pa_R		Pt_R		Sc_R		Ma_R	
	CS	RC	CS	RC	CS	RC	CS	RC	CS	RC	CS	RC
RCd		.37[H1a]		.39		.31		.32		.04		.16
1	.21	.15	.34[H1b]	.25[H1b]	.18	.16	.20	.16	.04	.03	.07	.08
2	.19[H1a,c]	.09[H1c]	.30	.21	.08	-.03	.21	.12	.10	-.08	-.06	-.18
3	.10	.20	.19	.29	.04	.25	.11	.20	-.02	.04	-.03	.12
4	.28	.18	.36	.29	.29	.20	.29	.19	.07	-.03	.18	.17
6	.16	-.03	.17	.03	.19[H1d]	.08[H1d]	.17	.06	-.01	-.14	.17	.12
7	.29	.22	.34	.25	.25	.23	.27[H1e]	.15[H1e]	.03	.02	.13	.16
8	.25	.17	.30	.13	.24	.19	.19	.09	-.02[H1f]	.05[H1f]	.18	.21
9	.09	.11	.05	.09	.14	.13	.10	.04	.00	-.04	.17[H1g]	.23[H1g]

Note. CS = Clinical Scale; RC = Restructured Clinical Scale; RCd = Demoralization; 1 = Hypochondriasis; RC1 = Somatic Complaints; 2 = Depression; RC2 = Low Positive Emotions; 3 = Hysteria; RC3 = Cynicism; 4 = Psychopathic Deviate; RC4 = Antisocial Behavior; 6 = Paranoia; RC6 = Ideas of Persecution; 7 = Psychasthenia; RC7 = Dysfunctional Negative Emotions; 8 = Schizophrenia; RC8 = Aberrant Experiences; 9 = Hypomania; RC9 = Hypomanic Activation; D_r = Depression; Hs_R= Hypochondriasis; Sc_R = Schizophrenia; Pt_R= Psychasthenia; Ma_R = Hypomania.
[Hxx] Correlations used to evaluate hypotheses are indicated by their specific hypothesis value.

than with Sc (rs = .05 and -.02, respectively, Steiger's z = 1.69, p > .05), as did Ma_R with RC9

than with Ma (rs = .23 and .17, respectively, Steiger's z = 0.98, p > .05), though these differences

were not significant. Similarly, D_R did show stronger correlation with D (r = .19) than with

RC2 (r = .09, Steiger's z = 1.73, p > .05), though this difference did not reach significance. In all

other SRF Scales, conceptually relevant SRF Scales showed stronger correlations with Clinical

than respective RC Scales including: Hs_R with Hs (r = .34) than with RC1 (r = .25, Steiger's z

= 3.68, p < .001), Pa_R with Pa (r = .19) than with RC6 (r =.08, Steiger's z = 1.92, p = .05), and

Pt_R with Pt (r = .27) than with RC7 (r = .15, Steiger's z = 3.06, p < .01). Hypothesis H1e

(RC7) was supported, and H1c (RC2), H1f (RC8), and H1g (RC9) followed hypothesized trends,

but were partially supported due to failure to meet either (a) magnitude of correlation difference

greater than .05, or (b) significance of difference in correlation strength indicated by Steiger's z

test. Hypotheses H1a (RCd), H1b (RC1), and H1d (RC6) were not supported.

Discriminant Validity: Korean Normative Couple Sample: Spouse Rating Form (SRF)

Similar to results of convergent validity, a majority of RC Scales showed poorer

discriminant validity of SRF Scales than respective Clinical Scales (Table 7 discriminant validity

calculations can be seen in Appendix E). RCd showed greater discriminant validity of D_R (r_{dis}

= .13) than did D (r_{dis} = .04), as did RC8 (r_{dis} = -.11) of Sc_R than did Sc (r_{dis}= -.25), and RC9 of

Ma_R (r_{dis} = .15) than did Ma (r_{dis} = .09). All other SRF Scales were better discriminated by

Clinical than RC Scale versions including D of D_R (r_{dis} = .04) than did RC2 (r_{dis} = -.03), Hs of

Hs_R (r_{dis} = .20) than did RC1 (r_{dis} = .13), and Pt of Pt_R (r_{dis} = .06) than did RC7 (r_{dis} = -.03).

Pa (r_{dis} = .05) showed slightly greater discriminant validity of Pa_R than did RC6 (r_{dis} = .00).

Hypotheses H2e (RC7), H2f (RC8), and H2g (RC9) were supported. Hypotheses H2a (RCd),

H2b (RC1), H2c (RC2), and H2d (RC6) were not supported.

Incremental Utility: Clinical Sample: Structured Clinical Interview for Diagnosis on DSM-IV (SCID)

Table 8 presents the results of the hierarchical logistic regression models predicting SCID diagnosis versus other diagnostic status. A Clinical Scale was entered in step one and the corresponding RC Scale was entered in step two. Results of models in which the RC Scale was entered in step one and the respective Clinical Scale was entered in step two are listed in the two columns on the far right of the table. In prediction of roughly half of SCID diagnoses, RC Scales contributed greater incremental utility to that of respective Clinical Scales than that of the reverse order of predictor entry. Incremental utility of RC8 to Sc ($\Delta R^2 = .062$) was greater than that of Sc to RC8 ($\Delta R^2 = .042$) in prediction of diagnosis of Schizophrenia with both increments being significant (panel six). RC9 also contributed significant and greater incremental utility to Ma ($\Delta R^2 = .066$) in prediction of diagnosis of Bipolar Disorder than that of Ma to RC9 ($\Delta R^2 = .000$) (panel seven). In the model predicting diagnosis of Schizophrenia, incremental utility of RC6 to Pa was greater than that of Pa to RC6 ($\Delta R^2 = .015$ vs. .000), though neither increment was significant (panel four). In the models predicting Major Depressive Disorder (MDD), D added significant and greater incremental utility to RCd ($\Delta R^2 = .046$) than that of RCd to D ($\Delta R^2 = .002$) (panel one). Hs contributed significant and greater incremental utility to RC1 ($\Delta R^2 = .020$) than that of RC1 to Hs ($\Delta R^2 = .004$) (panel two). D also added significant and greater incremental utility to RC2 ($\Delta R^2 = .034$) (panel three) than that of RC2 to D ($\Delta R^2 = .000$). In prediction of diagnosis of Anxiety Disorder, incremental utility of RC7 to Pt was negligible and similar to that of Pt to RC7 ($\Delta R^2 = .000$ and .001, respectively) (panel five). Hypotheses H3a (RCd), H3c (RC2), H3f (RC8), and H3g (RC9) were supported. Although results followed expected trends in comparative utility, respective RC and Clinical Scales contributed relatively similar and non-

significant proportions of incremental utility above that of the other, failing to support hypothesis

H3d (RC6). Hypotheses H3b (RC1) and H3e (RC7) were not supported.

Table 8. *Hierarchical Logistic Regression Analyses Results: Predicting SCID Diagnostic Status from Conceptually Relevant Restructured Clinical (RC) and Clinical Scales*

Step	Scale	β	OR	Model fit χ^2	$\Delta\chi^2$	R^2	ΔR^2	$\Delta\chi^2_1$	ΔR^2_1
MDD									
Step 1	D	.08**	1.08	21.22***		.081			
Step 2	RCd	.02	1.02	21.74***	.52	.083	.002	12.19***	.046
SOM									
Step 1	Hs	.15*	1.16	14.68***		.062			
Step 2	RC1	-.07	.93	15.67***	.99	.066	.004	4.86*	.020
MDD									
Step 1	D	.09**	1.09	21.22***		.081			
Step 2	RC2	.01	1.01	21.26***	.04	.081	.000	9.08**	.034
SCZ									
Step 1	Pa	-.00	1.00	5.94*		.024			
Step 2	RC6	.11	1.12	9.53**	3.59	.039	.015	.01	.000
ANX									
Step 1	Pt	-.02	.99	.73		.003			
Step 2	RC7	.01	1.01	.78	.05	.003	.000	.38	.001
SCZ									
Step 1	Sc	-.05**	.95	.06		.000			
Step 2	RC8	.21***	1.23	15.29***	15.23***	.062	.062	10.49**	.042
BP									
Step 1	Ma	.01	1.01	33.28***		.123			
Step 2	RC9	.16***	1.17	52.42***	19.14***	.189	.066	.02	.000

Note. β = standardized regression coefficients in the full model where all predictors are entered. OR=Odds Ratio (e^b). Nagelkerke R^2 was used. $\Delta\chi^2_1$ = Incremental χ^2 added by Clinical Scale over and above RCX. ΔR^2_1 = Incremental R^2 added by Clinical Scale over and above RCX. MDD = Major Depressive Disorder; SCZ = Schizophrenia; BP = Bipolar Disorder; RCd = Demoralization; RC2 = Low Positive Emotions; RC6 = Ideas of Persecution; RC8 = Aberrant Experiences; RC9 = Hypomanic Activation; D = Depression; Pa = Paranoia; Sc = Schizophrenia; Ma = Hypomania
* = $p < .05$. ** = $p < .01$. *** = $p < .001$.

Results of exploratory analyses examining the impact of adding the demoralization factor to each RC Scale are presented in Appendix F. In prediction of each criterion measure, RCd entered simultaneously with the RC Scale in step one, and the respective Clinical Scale entered in step two. The order of entry was then reversed in which the Clinical Scale entered in step one, and the simultaneous entry of RCd with the RC Scale entered in step two. In prediction of nearly all SCID diagnostic categories, the simultaneous entry of RCd with the RC Scale contributed greater incremental utility to that of the respective Clinical Scale than that of the respective Clinical Scale to RCd with the corresponding RC Scale. Examples included RC1 with RCd to Hs than that of Hs to RC1 with RCd (ΔR^2 = .089 and .032, respectively) in prediction of diagnosis of Somatoform Disorder (panel two), RC6 with RCd to Pa than that of Pa to RC6 with RCd (ΔR^2 = .084 and .011, respectively) in prediction of diagnosis of Schizophrenia (panel three), RC7 with RCd to Pt than that of Pt to RC7 with RCd (ΔR^2 = .062 and .019, respectively) in prediction of diagnosis of Anxiety Disorder (panel four), RCd with RC8 in prediction of diagnosis of Schizophrenia to Sc than that of Sc to RC8 with RCd (ΔR^2 = .073 and .002, respectively) (panel five), and RCd with RC9 to Ma than that of Ma to RCd with RC9 (ΔR^2 = .067 and .001, respectively) (panel six). In contrast, incremental utility of D to RC2 with RCd (ΔR^2 = .025) was still greater than that of RC2 with RCd to D (ΔR^2 = .002) in prediction of diagnosis of Major Depressive Disorder (MDD) (panel one).

Incremental Utility: Clinical Sample: Patient Description Form (PDF)

Results of the hierarchical regression analyses predicting Patient Description Form (PDF) Scales are presented in Table 9. Models in which a Clinical Scale was entered in step one and respective RC Scale was entered in step two are shown in the left portion of the table, while the

far right column presents results of models in which the RC Scale was entered in step one, and

the respective Clinical Scale was entered in step two.

Table 9. *Hierarchical Regression Analysis Results: Predicting Patient Description Form (PDF) Scales from Restructured Clinical (RC) and Clinical Scales*

Step	Scale	R^2	ΔR^2	β	r	sr^2_1
Depressed						
Step 1	D	.271***		.31***	.52***	
Step 2	RCd	.382***	.111***	.40***	.56***	.067***
Somatic						
Step 1	Hs	.231***		.96***	.48***	
Step 2	RC1	.261***	.030***	-.51***	.39***	.106***
Depressed						
Step 1	D	.271***		.40***	.52***	
Step 2	RC2	.283***	.012*	.16*	.46***	.068***
Suspicious						
Step 1	Pa	.106***		.16*	.33***	
Step 2	RC6	.123***	.017**	.21**	.34***	.010*
Anxiety						
Step 1	Pt	.066***		.34**	.26***	
Step 2	RC7	.068***	.002	-.09	.20***	.028**
Insecure						
Step 1	Pt	.126***		.44***	.35***	
Step 2	RC7	.128***	.002	-.10	.28***	.048***
Obs-Comp						
Step 1	Pt	.015*		.22*	.12*	
Step 2	RC7	.018*	.003	-.11	.08	.012*
Psychotic						
Step 1	Sc	.008		-.32***	.09	
Step 2	RC8	.115***	.107***	.52***	.28***	.039***
Narcissistic						
Step 1	Ma	.037***		.12	.19***	
Step 2	RC9	.041***	.004	.10	.19***	.005
Agitated						
Step 1	Ma	.029**		.01	.17**	
Step 2	RC9	.045***	.016*	.21*	.21***	.000

Note. β = standardized regression coefficients in the full model where all predictors are entered. r = zero-order correlation of each predictor with a criterion. sr^2_1 = Incremental R^2 added by Clinical Scale X over and above RCX. RC = Restructured Clinical Scale. RCd = Demoralization; RC1 = Somatic Complaints; RC2 = Low Positive Emotions; RC6 = Ideas of Persecution; RC7 = Dysfunctional Negative Emotions; RC8 = Aberrant Experiences; RC9 = Hypomanic Activation; Hs = Hypochondriasis; D = Depression; Pa = Paranoia; Pt = Psychasthenia; Sc = Schizophrenia; Ma = Hypomania; Obs-Comp = Obsessive-Compulsive.
* = $p<.05$. ** = $p<.01$. *** = $p<.001$.

In prediction of half of PDF Scales, incremental utility of RC Scales to respective Clinical Scales was similar to greater than that of the Clinical Scale to the corresponding RC Scale. For example, in prediction of Depressed, incremental utility of RCd to D ($\Delta R^2 = .111$) was greater than that of D to RCd ($\Delta R^2 = .067$) (panel one), as was incremental utility of RC6 to Pa ($\Delta R^2 = .017$) than that of Pa to RC6 ($\Delta R^2 = .010$) in prediction of Suspicious (panel four), RC8 to Sc in prediction of Psychotic ($\Delta R^2 = .107$) than that of Sc to RC8 ($\Delta R^2 = .039$) (panel eight), and RC9 to Ma ($\Delta R^2 = .016$) than that of Ma to RC9 ($\Delta R^2 = .000$) in prediction of Agitated (panel ten) with all increments being significant. In prediction of Narcissistic, incremental utility of Ma to RC9 and that of RC9 to Ma were both negligible ($\Delta R^2 = .005$ and $.004$, respectively).

In prediction of all other PDF Scales, incremental utility of each Clinical Scale to respective RC Scale was greater than that of the RC Scale to the corresponding Clinical Scale. Examples included incremental utility of Hs to RC1 ($\Delta R^2 = .106$) than that of RC1 to Hs ($\Delta R^2 = .030$) in prediction of Somatic Complaints and D to RC2 ($\Delta R^2 = .068$) than that of RC2 to D ($\Delta R^2 = .012$) in prediction of Depressed (third panel) with all increments being significant. Incremental utility of Pt to RC7 in prediction of Anxiety, Insecure, and Obsessive-compulsive ($\Delta R^2 = .028$, $.048$, and $.012$, respectively) was significant and greater than that of RC7 to Pt ($\Delta R^2 = .002$, $.002$, and $.003$, respectively). Hypotheses H3c (RC2), H3d (RC6), H3e (RC7), and H3f (RC8) were supported, and Hypothesis H3g (RC9) was supported in one of two criterion measures. Hypotheses H3a (RCd), and H3b (RC1) were not supported.

Results of exploratory analyses with RCd entered simultaneously with the RC Scale in prediction of conceptually relevant PDF Scales are presented in Appendix G. Incremental utility of the simultaneous entry of RCd with the conceptually relevant RC Scale was greater than that of the respective Clinical Scale to RCd with the corresponding RC Scale in prediction of a

majority of PDF Scales. This included the model predicting Depressed ($\Delta R^2 = .118$ vs. .017) (panel one), Suspicious ($\Delta R^2 = .035$ vs. .023) (panel three), Insecure ($\Delta R^2 = .047$ vs. .000) (panel five), Obsessive-Compulsive (panel six) ($\Delta R^2 = .012$ vs. .001), Psychotic ($\Delta R^2 = .124$ vs. .001) (panel 7), Narcissistic ($\Delta R^2 = .025$ vs. .007) (panel eight), and Agitated ($\Delta R^2 = .018$ vs. .000) (panel nine). In contrast, incremental utility of Hs to RC1 with RCd ($\Delta R^2 = .122$) was greater than that of RC1 with RCd to Hs ($\Delta R^2 = .058$) (panel 2). Similarly, incremental utility of Pt to RC7 with RCd was greater than that of RC7 with RCd to Pt ($\Delta R^2 = .018$ and .002, respectively) in the model predicting Anxious (panel four).

Incremental Utility: Korean Normative Couple Sample: Spouse Rating Form (SRF)

In prediction of just under half of Spouse Rating Form (SRF) Scales (Table 10), results of hierarchical regression analyses showed incremental utility of RC Scale to respective Clinical Scale was similar to greater than that of the RC Scale to the corresponding Clinical Scale. Incremental utility of RCd to D was significant and greater than that of D to RCd in prediction of D_R (panel one) ($\Delta R^2 = .104$ and .000, respectively), as was RC9 to Ma ($\Delta R^2 = .024$) than that of Ma to RC9 ($\Delta R^2 = .001$) in prediction of Ma_R (panel seven). As shown in panel six, incremental utility of RC8 to Sc in prediction of Sc_R was similar to that of Sc to RC8 ($\Delta R^2 = .012$ and .010, respectively) with neither increment being significant. Incremental utility of Clinical Scales to that of respective RC Scales was greater than that of the RC Scale to the corresponding Clinical Scale in all other models. Incremental utility of D to RC2 ($\Delta R^2 = .029$) in prediction of D_R was significant and greater than that of RC2 to D ($\Delta R^2 = .002$) (panel two), as was Hs to RC1 than that of RC1 to Hs ($\Delta R^2 = .086$ and .033, respectively) in prediction of Hs_R (panel three), Pa to RC6 ($\Delta R^2 = .031$) than that of RC6 to Pa ($\Delta R^2 = .004$) in prediction of Pa_R (panel four), and Pt to RC7 ($\Delta R^2 = .068$) than that of RC7 to Pt ($\Delta R^2 = .016$) in prediction of

56

Pt_R (panel five). Hypotheses H3c (RC2), H3e (RC7), and H3g (RC9) were supported.

Hypotheses H3a (RCd), H3b (RC1), H3d (RC6), and H3f (RC8) were not supported.

Table 10. *Hierarchical Regression Analysis Results: Predicting Spouse Rating Form (SRF)*
Scales from Conceptually Relevant Restructured Clinical (RC) and Clinical Scales

Step	Scale	R^2	ΔR^2	β	r	sr^2_1
D_R						
Step 1	D	.035**		.02	.19**	
Step 2	RCd	.139***	.104***	.36***	.37***	.000
Hs_R						
Step 1	Hs	.113***		.80***	.34***	
Step 2	RC1	.146***	.033**	-.50**	.25***	.086***
D_R						
Step 1	D	.035**		.22*	.19**	
Step 2	RC2	.037*	.002	-.05	.09	.029*
Pa_R						
Step 1	Pa	.034**		.25**	.19**	
Step 2	RC6	.038*	.004	-.09	.08	.031**
Pt_R						
Step 1	Pt	.074***		.45***	.27***	
Step 2	RC7	.090***	.016	-.22	.15*	.068***
Sc_R						
Step 1	Sc	.001		-.17	-.02	
Step 2	RC8	.013	.012	.18	.05	.010
Ma_R						
Step 1	Ma	.030*		.05	.17*	
Step 2	RC9	.054**	.024*	.20*	.23**	.001

Note. β = standardized regression coefficients in the full model where all predictors are entered. r = zero-order correlation of each predictor with a criterion. sr^2_1 = Incremental R^2 added by Clinical Scale X over and above RCX (and RCd). RC = Restructured Clinical Scale. RCd = Demoralization; RC1 = Somatic Complaints; RC2 = Low Positive Emotions; RC6 = Ideas of Persecution; RC7 = Dysfunctional Negative Emotions; RC8 = Aberrant Experiences; RC9 = Hypomanic Activation; Hs = Hypochondriasis; D = Depression; Pa = Paranoia; Pt = Psychasthenia; Sc = Schizophrenia; Ma = Hypomania; D_r = Depression; Hs_R= Hypochondriasis; Sc_R = Schizophrenia; Pt_R= Psychasthenia; Ma_R = Hypomania.
*= $p<.05$. **= $p<.01$. ***= $p<.001$.

Incremental utility of the simultaneous entry of the conceptually relevant RC Scale with RCd to Clinical Scale was similar to greater than that of respective Clinical Scales to the simultaneous entry of RCd with the RC Scale version in prediction of all SRF Scales (see Appendix H). Comparisons of incremental proportions are as follows: RC2 with RCd to D (ΔR^2 = .104) than that of D to RC2 with RCd (ΔR^2 = .000) in prediction of D_R (panel one), RC1 with RCd to Hs (ΔR^2 = .102) than that of Hs to RC1 with RCd (ΔR^2 = .059) in prediction of Hs_R (panel two), RC6 with RCd to Pa (ΔR^2 = .073) than that P to RC6 with RCd (ΔR^2 = .009) (panel three) in prediction of Pa_r, RC7 with RCd to Pt (ΔR^2 = .051) than that Pt to RC7 with RCd (ΔR^2 = .005) (panel four), and RC9 with RCd to Ma in prediction of Ma_R (ΔR^2 = .027) than that of Ma to RC9 with RCd (ΔR^2 = .001); however, neither increment reached significance (panel six). The combined entry of RC8 with RCd contributed roughly similar proportion of incremental utility to Sc than that of Sc to RC8 with RCd (ΔR^2 = .023 and .021, respectively) in prediction of Sc_R (panel five).

CHAPTER V

DISCUSSION

Current study compared the utility of the Korean MMPI-2 Restructured Clinical (RC) Scales with that of the Clinical Scales. With a small degree of disagreement across the three criterion measures (SCID diagnoses, PDF Scales, and SRF Scales), results of the study were consistent with previous research indicating a majority of RC Scales do provide equivalent to modest improvements in utility above that of their Clinical Scale version while containing substantially reduced item counts (Sellbom & Ben-Porath, 2005; Sellbom, Ben-Porath, et al., 2006; Sellbom, Graham, et al., 2006; Simms et al., 2005). More specifically, RC Scales yielded similar to greater convergent validity of a majority of SCID diagnoses, half of conceptually relevant Patient Description Form (PDF) Scales, and half of SRF Scales than Clinical Scale versions. In comparisons of discriminant validity, results offered more substantial support, as RC Scales showed similar to greater discriminant validity of most SCID diagnoses and PDF Scales, as well as roughly half of SRF Scales. With the exception of Anxiety Disorder, all other hypotheses of convergent and discriminant validity of SCID diagnoses were partially or fully supported. Roughly half of hypotheses of convergent and discriminant validity of PDF and SRF Scales were partially or completely supported.

Results of hierarchical logistic and hierarchical regression analyses showed that RC Scales contributed similar to greater proportion of incremental utility to respective Clinical Scales than that of the Clinical Scale to the RC Scale version in prediction of roughly half of SCID diagnostic categories and PDF Scales, and just under half of SRF Scales. Roughly half of models predicting SCID diagnosis and PDF Scales followed hypothesized trends (H3a-H3g); however, just under half of the models predicting SRF Scales supported study hypotheses.

59

Convergent and Discriminant Validity

Clinical Sample: Structured Clinical Interview for Diagnosis on DSM-IV (SCID)

Convergent and discriminant validities of RC Scales of a majority of conceptually relevant SCID diagnoses, compared to Clinical Scales, were similar to slightly greater. With few exceptions, results supported or partially supported hypothesized trends of convergent and discriminant validity. For example, RC6, RC8, and RC9 showed slightly greater to greater, and RC1 showed similar convergent and discriminant validity of conceptually relevant SCID diagnoses as did respective Clinical Scales. As hypothesized, RCd and RC2 showed slightly weaker convergent and relatively similar discriminant validity of diagnosis of MDD than did D. Surprisingly, comparisons of RC7 to Pt did not follow hypothesized trends as both scales showed similar convergent and discriminant validity of diagnosis of Anxiety Disorder.

Unexpectedly, RC and Clinical Scales exhibited much weaker correlations with conceptually relevant SCID diagnoses than reported in previous studies using SCID diagnoses as criterion (Simms et al., 2005). For example, most correlations in the current study were relatively weak ($rs < .20$); Simms and colleagues reported mostly moderate to strong correlations. Weaker correlations displayed in the current study led to restricted range in comparing RC and Clinical Scales. Although most comparisons following hypothesized trends, differences between correlations of RC and Clinical Scales with conceptually relevant SCID diagnoses were not large enough to meet study criterion. As a result, many study hypotheses were only partially instead of fully supported.

Weakness of correlation strength was especially evident in the comparison of convergent validity of diagnosis of Anxiety Disorder of RC7 with Pt (the only hypotheses of SCID diagnosis not to be fully or partially supported). Correlations of diagnosis of Anxiety Disorder with RC7

and Pt indicated much weaker relationships in the current study ($rs = -.03$ and $-.04$, respectively) than those reported by Simms et al. (2005) ($rs = .32$ and $.37$, respectively). In addition, comparison of discriminant validity of RC7 with that of Pt indicated that both scales were more strongly related to conceptually unrelated SCID diagnoses to a similar extent (equivalent negative discriminant validity values). Evidence of poor discriminant validity has been reported in several past studies. For instance, Simms and colleagues also reported poor discriminant validity of RC7 due to strong correlations with criterion measures related to depression. In addition, Sellbom, Ben-Porath, and colleagues (2006) write RC7 captures symptoms of negative emotionality, a broader trait which also includes the construct of anxiety, along with other symptoms of alienation (e.g., paranoia/mistrust). Simms et al. argue further revision on item content of RC7 should be conducted to strengthen its convergent and discriminant validity of anxiety related disorders above that of the current items. Current study findings appear to reinforce this argument.

Aside from the specific example of RC7, differences in sample characteristics may explain overall disparity in correlation strength between studies. Simms and colleagues (2005) conducted research using military veterans, whereas sample of the current study comprised clinical patients. Due to traumatic events experienced in the military, veterans may have reported complaints across a greater range of symptoms of psychological maladjustment (e.g., cognitive dysfunction and chronic widespread pain), as well as across a wider degree of symptom severity than did clinical patients. Greater variance in response styles would lead to increased correlation strength resulting in increased differences in range of correlations reported in each study.

Differences in SCID diagnoses presented in each sample may have further altered validity comparisons, especially discriminant validity. While both studies included diagnoses of

Major Depressive Disorder (MDD), Somatoform, and Anxiety Disorders, current sample patients comprised diagnoses of Bipolar Disorder and Schizophrenia, while Simms and colleagues (2005) reported diagnosis of Substance Abuse. Since discriminant validity is a function of correlation strength with both related and unrelated criterion measures, findings may differ somewhat among studies sampling from different diagnostic categories. For example, Simms and colleagues reported convergent validity of Substance Abuse was the only diagnostic category in which the conceptually relevant RC Scale outperformed that of the related Clinical Scale. Since relationships of Substance Abuse with RC and Clinical Scales differed from those of other SCID diagnoses, inclusion of this diagnostic category may have altered study results in comparing utility (especially discriminant validity) among RC and Clinical Scales.

Finally, proportion of individuals diagnosed with co-morbid disorders may have weakened correlations in the current study as compared to those reported by Simms and colleagues (2005). With the relatively large proportion (26%) of diagnostic co-morbidity (see Appendix A) observed in the current study, symptoms reported by individuals diagnosed with specific disorders may not be as mutually distinct as would be expected. This non-distinct reporting of symptoms across diagnostic categories may have decreased strength of relationships of RC and Clinical Scales with conceptually relevant and increased relationship strength with conceptually unrelated SCID diagnoses, thereby diminishing correlation magnitude differences as compared to those reported in Simms et al. (2005).

Clinical Sample: Patient Description Form (PDF)

RC Scales showed similar to greater convergent validity of just over half of PDF Scales and similar to greater discriminant validity of a majority of PDF Scales than did respective Clinical Scales. Results supported or partially supported just over half of study hypotheses of

convergent validity and just under half of hypotheses of discriminant validity. Despite failing to support large number of hypotheses, several comparisons did nevertheless follow findings from past research using multiple external criterion measures of psychological symptoms (Sellbom, Ben-Porath, et al., 2006). For example, as compared to respective Clinical Scales, RC9 showed slightly greater to greater convergent and discriminant validity of conceptually relevant PDF scales, and RC2 showed weaker convergent and relatively similar discriminant validity. In addition, RC7 showed weaker convergent validity than did Pt. In terms of correlation magnitude, both studies reported small to moderate correlations of RC and Clinical Scales with conceptually relevant external criterion measures. Correlations of RC9 reported by Sellbom, Ben-Porath et al. with conceptually relevant SCL-90-R Hostility and CDF Scale Aggressive ($rs =.16$ and .20, respectively) were nearly identical to findings of the current study of RC9 with PDF Narcissistic ($r =.19$) and Agitated ($r =.21$), as well as correlation of RC7 with SCL-90-R Anxiety ($r =.21$) to that of the current study with PDF Scale Anxious ($r =.20$).

Comparing results of convergent and discriminant validity of RC1 to that of Hs showed a curious difference between current and previous study results. Sellbom, Ben-Porath, and colleagues (2006) reported RC1 and Hs to yield similar convergent validity of multiple criterion measures (range $rs = .25$ to .30) yielding correlations of small to moderate strength. They also reported equivalent discriminant validities for each scale (range $rs_{dis} = .13$ to .17). In contrast, the current study showed Hs to display greater convergent ($rs = .48$ and .39, respectively) of conceptually relevant PDF Somatic Complaints Scale than did RC1 with moderate to strong correlations. Current results also showed Hs yielded greater discriminant validity of Somatic Complaints ($rs_{dis} = .34$ and .27, respectively) than did RC1. An explanation for this finding may be found in the tendency of Korean individuals to typically express psychological distress as

somatic symptoms in place of frustration or anger (Pang, 1990). Since Koreans more frequently report somatic complaints than do Americans, reduction in both item count, as well as symptom heterogeneity, may have more severely weakened convergent and discriminant validity of RC1 over that of Hs in Korean than in U.S. samples. For example, several items included in Hs are also included on scales measuring symptoms of depression, such as D (e.g., i175: I feel weak all over much of the time, and i39: My sleep is fitful and disturbed). This creates a high degree of item overlap with other symptoms of psychological maladjustment, which has been mentioned in past research (e.g., anxiety and depressive disorders; Sellbom, Ben-Porath, et al., 2006; Simms et al., 2005). Perhaps this greater degree of item overlap present in Hs provides a more complete assessment of somatic complaints in Korean individuals than does RC1. Therefore, use of symptom heterogeneous instead of symptom homogeneous scales may be necessary to accurately assess somatic complaints in Korean individuals.

The failure of RC6 to provide both greater convergent and discriminant validity of conceptually relevant scales than Pa was another comparison that failed to follow hypothesized trends. Current study results showed RC6 to yield similar convergent validity and only slightly greater discriminant validity than did Pa, whereas previous research (Sellbom, Ben-Porath, et al., 2006) reported greater convergent and discriminant validity for RC6 above that of Pa. There are several possible explanations for this difference. First, although most correlations were of similar strength between the two studies, current study reported moderate correlations of both Pa and RC6 with predicted PDF Scale Suspicious; Sellbom, Ben-Porath, and colleagues reported only weak correlations. Differences in magnitude of correlation strength may result from differences in sample types. Clinical inpatients and outpatients diagnosed with one or more SCID diagnoses may be experiencing psychological maladjustment to a more extreme level, leading to more

64

frequent and more widespread endorsement of symptoms of distrust and paranoia than do clients

of a psychology college clinic, the majority of participants included by Sellbom, Ben-Porath, and

colleagues. In fact, authors reported symptoms of Paranoia to be especially restricted in range

(highly skewed) due to low sample base rates. With more widespread symptoms being reported

in the current study, naturally, the broader item content of Pa, over that of RC6, would result in

better utility. In addition, differences in measures selected to assess Paranoia may have

increased result discrepancies between studies. Sellbom, Ben-Porath, et al. report their choice of

Suspiciousness may have been a poor measure of utility of Pa and RC6 due to its close

relationship with symptoms of alienation and negative emotionality rather than with symptoms

of paranoia.

Another unpredicted finding was that RCd showed slightly greater convergent and

greater discriminant validity of conceptually relevant PDF Scale than did D in current study,

whereas Sellbom, Ben-Porath, and colleagues (2006) showed RCd displayed similar to slightly

weaker convergent and discriminant validity of most conceptually relevant criterion measures.

Although results did not follow hypothesized trends, differences in utility of RCd and D were

mostly small and non-significant indicated by Steiger's *z*. Patterns in comparative utility may not

be large enough to indicate meaningful support for utility of either RCd or D above that of the

other. Interestingly, these results compared more closely with those reported in studies using

self-report criterion measures (e.g., Sellbom & Ben-Porath, 2005; Sellbom, Graham, et al., 2006)

in which RCd provided similar to greater convergent validity over that of D. Considering

evidence from research utilizing both external and self-report measures, current study results

provide moderate support for the success RCd as able to provide relatively similar convergent

and discriminant validity as that of D, despite severely decreased item count and content.

Although RC7 and Pt did conform to expected comparisons of convergent validity, Pt did not provide greater discriminant validity of conceptually relevant PDF Scales than did RC7; both scales provided similar discriminant validity. This finding may be explained in part by research suggesting that discriminant validity, over that of convergent validity, has been improved in the revision process to the RC Scales (e.g., Sellbom & Ben-Porath, 2005; Sellbom, Ben-Porath, et al., 2006; Sellbom, Ben-Porath, McNulty, Arbisi, & Graham, 2006). Discriminant validity of RC7, as compared to that of Pt, appears to have been improved, at least within PDF Scales. As this finding was not observed in discussion of SCID diagnoses, additional research, including the specific constructs captured by this scale, is most likely required as performance across criterion measures tends to be sporadic.

Korean Normative Couple Sample: Spouse Rating Form (SRF)

RC Scales showed similar to slightly greater convergent validity of roughly half of SRF Scales and similar to slightly greater discriminant validity of just shy of half of SRF Scales with half of comparisons failing to follow hypothesized trends in comparative utility. Many of the findings that were inconsistent with study hypotheses were also noted in results of PDF Scales. For instance, RCd provided slightly greater convergent and discriminant validity than the respective Clinical Scale, while utility of RC1 and RC6 was weaker than expected. Only comparisons of convergent and discriminant validity of RC7, RC8, and RC9 with that of respective Clinical Scales followed study hypotheses.

Despite failure to conform to study hypotheses, magnitude of correlation strength was similar to that reported by previous research (e.g., Sellbom, Ben-Porath, et al., 2005). For example, correlations of both current and past studies yielded small to medium correlations of

conceptually relevant criterion with RC9, while both studies showed correlations of RCd with conceptually relevant measures to be medium.

The finding that Pt showed both greater convergent (H1e) and discriminant validity (H2e) of Pt_R than did RC7 was one of the more surprising results supporting study hypotheses. Results of SCID diagnoses failed to support either hypotheses, while results of PDF Scales supported only convergent (H1e), but not discriminant validity hypotheses (H2e). As mentioned previously, performance of RC7 appears to be inconsistent. Clearly, further research, as well as item content revision, is needed to improve utility and dependability of this scale across varying sample types and criterion measures.

<div align="center">Incremental Utility</div>

Clinical Sample: Structured Clinical Interview for Diagnosis on DSM-IV (SCID)

In prediction of roughly half of SCID diagnoses, RC Scales contributed similar to greater incremental utility to respective Clinical Scales than that of the Clinical Scale to the corresponding RC Scale. Results of just over half of the models predicting diagnostic status followed hypothesized trends of incremental utility including those with predictors of RCd, RC2, RC8, and RC9, while RC1, RC6, and RC7 did not conform to study hypotheses.

Results did resemble those reported in previous research using SCID diagnoses as criterion measures (Simms et al., 2005). Both studies showed proportions of incremental utility to be mostly small and non-significant. For example, incremental utility of the hypothesized RC Scale to the respective Clinical Scale reached significance in only two of the seven models, whereas proportions of incremental utility of the Clinical to the respective RC Scale reached significance in four models in the current study. Simms and colleagues reported neither RC nor

Clinical Scales contributed significant incremental utility to that of the other with the exception of the model predicting diagnosis of substance abuse in which RC Scales contributed significant and greater proportions of incremental utility to Clinical Scales, while Clinical Scales contributed no significant incremental utility to RC Scales. Similarities in findings remained despite differences in method of predictor entry. Simms and colleagues compared proportion of incremental utility contributed by entering the eight Clinical Scales as a block in one step and the corresponding eight RC Scales as a block in a separate step. In comparison, current study entered only one RC Scale in one step and its respective Clinical Scale in a separate step. For example, in prediction of diagnosis of Somatoform and Anxiety Disorders, Simms et al. (2005) reported RC and Clinical Scales contributed similar and relatively negligible proportions of incremental utility to that of each other ($\Delta R^2 < .015$). These results were very similar to those of the current study with reported proportions $\leq .020$, with the exception that proportion contributed by Hs to RC1 did reach significance in the current study. Current study findings support Simms and colleagues' statement of which use of binary criterion measures weakens results.

One interesting comparison that could be made between studies was seen in examining models predicting diagnosis of Major Depressive Disorder (MDD). Simms et al. (2005) showed RC Scales contributed similar proportion of incremental utility to Clinical Scales in prediction of both current and lifetime diagnosis of Depressive Disorders as Clinical Scales to RC Scales, while the current study showed D contributed significant and greater incremental utility to both RCd and to RC2 than that of either RCd to D or RC2 to D. Research utilizing methods of predictor entry similar to the current study (e.g., Sellbom, Graham, et al., 2006) reported both RCd and RC2 contributed smaller proportions of incremental utility to D than that of the reverse,

though this research used self-report criterion measures designed to correspond to SCID diagnoses.

Clinical Sample: Patient Description Form (PDF)

Results of hierarchical regression analyses showed RC Scales contributed similar to greater proportions of incremental utility to Clinical Scales than that of the Clinical Scale to the respective RC Scale in prediction of half of PDF Scales. In contrast to models predicting diagnostic status, RC and Clinical Scales contributed significant proportions of incremental utility above that of the other in a majority of models predicting PDF Scales. Just over half of study hypotheses (H3a-g) were supported.

Results found many similarities to those reported by Sellbom, Graham, et al. (2006) in which self-report criterion measure was used (Multiaxial Diagnostic Inventory; MDI). For example, current study showed RC8 and RC9 (in one of two PDF scales) contributed greater, while RC2 and RC7 contributed smaller proportions of incremental utility to respective Clinical Scales than that of the reverse order of predictor entry in prediction of all conceptually relevant criterion measures.

Despite similarities, several differences existed between studies. Current study found RC1 contributed a smaller proportion of incremental utility to Hs in prediction of conceptually relevant PDF Scale; Sellbom, Graham, et al. (2006) reported opposite patterns of results in prediction of both conceptually relevant MDI Factor and Diagnostic Scales. As current study results reported Hs also provided greater incremental utility to RC1 of prediction of diagnosis Somatoform Disorder than RC1 to Hs, as well as greater convergent and discriminant validity of PDF Scales than RC1, these findings provide additional evidence that the construct of somatic complaints is not being captured as successfully in the RC as compared to that of the original

69

Clinical Scale version. Mentioned previously, cultural differences in expression of psychological complaints between Korean and American cultures (Pang, 1990) are most likely a source of this failure to follow trends reported in U.S. MMPI-2 research.

In addition, current study showed RCd contributed greater proportion of incremental utility to D than that of D to RCd in prediction of conceptually relevant PDF Scale. These results differed from results reported by Sellbom, Graham, and colleagues (2006) reporting slightly smaller (.02) proportion of incremental utility contributed by RCd to D than that of D to RCd in prediction of both conceptually relevant MDI Factor and Diagnostic Scales. Although this difference was minimal and possibly not a stable finding to generalize to other samples, current study results do give some evidence supporting the ability of RCd to provide at least equivalent, if not slightly improved, utility despite reduced item complexity and item count as that of the much longer and symptom heterogeneous D.

Korean Normative Couple Sample: Spouse Rating Form (SRF)

In prediction of Spouse Rating Form (SRF) Scales, RC Scales contributed significant and greater proportion of incremental utility to Clinical Scales in prediction of just under half of the models predicting SRF Scales, while Clinical Scales contributed significant proportions of incremental utility to respective RC Scales in prediction of just over half of SRF Scales. Just under half of study hypotheses were supported. For example, comparisons of RC2, RC7, and RC9 with respective Clinical Scales followed expected patterns.

Results that did not follow predicted trends were relatively similar to those reported for PDF Scales in which comparisons of RCd and RC1 with Clinical Scale versions failed to support hypothesized predictions. However, RC6 and RC8 also failed to provide greater incremental utility to respective Clinical Scales than that of the Clinical Scale to RC6 or RC8. These findings

differed from models predicting PDF Scales. Differences in sample (clinical vs. normal) may explain some of these discrepancies in which reporting symptoms of suspiciousness, persecution, and unusual perceptions may not be as frequently endorsed in a normal as compared to a clinical sample. Thus, restricted range due to reduced frequency in item endorsement may explain why support for these two hypotheses (H3d and H3f) differed between criterion measures.

One general explanation for the differences found between the current and previous studies in comparative utility of RC to that of respective Clinical Scales may reflect the use of multiple (and diverse) criterion measures. Finding similar results across *all* criterion measures was highly unlikely, with many possibilities for exceptions to hypotheses to occur. For example, even Sellbom, Ben-Porath, et al. (2006) failed to report consistent patterns across multiple external criterion measures, especially for RCd, RC2, and RC7.

<center>Exploratory Analyses</center>

Results of exploratory analyses showed entering the factor of demoralization (RCd) simultaneously with each RC Scale did result in significant and greater incremental utility to the respective Clinical Scale than that of the Clinical Scale to the RC Scale version entered with RCd in prediction of nearly all hypothesized models. Averaged across all models predicting the three criterion measures, entering RCd simultaneously with the conceptually relevant RC Scale contributed roughly three times the proportion of incremental utility to the respective Clinical Scale ($\Delta R^2_{mean} = .057$) than that of the Clinical Scale to RCd with the RC Scale version ($\Delta R^2_{mean} = .018$).

For example, when RCd was entered into the model simultaneously with the hypothesized RC Scale as predictor of SCID diagnostic status, the combined entry of the respective RC Scale with RCd contributed significant incremental utility to the Clinical Scale in

<center>71</center>

all models except in the model predicting diagnosis of Major Depressive Disorder (MDD).

Proportion of contributed incremental utility from the simultaneous entry of RCd and the

hypothesized RC Scale to the respective Clinical Scale, on average, was greater than that of the

Clinical Scale to RCd with the RC Scale version (ΔR^2_{mean} = .063 vs. .015, respectively). The only

model in which this was the case was in the model predicting diagnosis of MDD in which D

contributed a greater (and significant) proportion of incremental utility to that of both RCd and

RC2 than that of the reverse order of entry (ΔR^2 = .025 and .002, respectively).

Results were relatively similar in the models predicting PDF Scales and SRF Scales.

Entering RCd simultaneously with the hypothesized RC Scale contributed significant and greater

proportion of incremental utility to the respective Clinical Scale, on average, as compared to that

of the Clinical Scale to RCd with the RC Scale version in prediction of PDF Scales (ΔR^2_{mean} =

.049 vs. .021, respectively) and SRF Scales (ΔR^2_{mean} = .063 vs. .016, respectively). The

simultaneous entry of the hypothesized RC Scale with RCd contributed greater proportion of

incremental utility to the respective Clinical Scale than that of the Clinical Scale to RCd with the

RC Scale version in prediction of seven out of nine PDF Scales, the exceptions being in models

predicting Somatic Complaints and Anxious. In models predicting SRF Scales, the simultaneous

entry of RCd with the hypothesized RC Scale contributed greater proportion of incremental

utility to the Clinical Scale than the Clinical Scale to RCd with the RC Scale version in

prediction of all SRF Scales with the exception of Sc_R, in which RCd with RC8 contributed

relatively similar incremental utility to Sc as that of Sc to RCd with RC8.

Within each of the three criterion measures, there were instances of suppression, which

may have altered results. However, these instances are most likely sample specific and should

not be taken as generalizable to other studies. For example, in the model predicting PDF Scale

Somatic Complaints, there was evidence of possible negative suppression: Hs may have acted as a negative suppressor variable, enhancing the importance of RC1 shown in the opposite signs of the standardized regression coefficient compared to correlation coefficient for RC1. In the model predicting SRF Scale Hs_R, evidence of negative suppression was noted indicating Hs may have been a negative suppressor variable increasing the predictive utility of RC1 by suppressing irrelevant variance.

In results of exploratory analyses, several models also showed possible example of suppression. For example, there was some evidence of suppression in models predicting diagnostic status. Classical suppression was noted in the model predicting Somatoform Disorder in which either Hs or RC1 may have acted as a suppressor variable suppressing irrelevant variance in RCd to increase its predictive utility (Appendix F, panel two). Classical suppression was also observed in the model predicting Schizophrenia: Pa or RC6 may have acted as suppressor variables to enhance the predictive utility of RCd (panel three). Negative suppression was noted in the model predicting Anxiety Disorder in which RC7 or RCd may have acted as a negative suppressor variable of Pt. These instances, however, are most likely sample specific and should not be taken as generalizable to other studies.

Strengths and Limitations

One of the key strengths of the current study was its ability to obtain and compare responses from both Korean psychiatric hospital patients as well as Korean adults from the general population. Previous studies have included comparisons from only one sample type (e.g., clinical: Sellbom, Graham, et al., 2006; military veteran: Simms et al., 2005), but rarely are studies able to compare results across both clinical and normal samples, as in the current study.

The fact that several of the findings provided consistent results across these two sample types strengthened confidence of study findings to be generalized across multiple application settings.

An additional strength of the current study is the quality of the clinical sample. The clinical sample comprised five SCID diagnostic categories containing relatively equal proportions of patients from each diagnosis. Previous studies have not been able to provide responses from this many diagnoses at one time or to offer such well balanced diagnostic categories (e.g., Simms et al., 2005). Therapists can therefore use this information to apply to more diverse clients. Diagnoses of the current study were completed by trained therapists over a relatively short period of time (less than two days) as compared to that of previous research which used student therapists with ratings completed over 21 days (Sellbom, Ben-Porath, et al., 2006). By completing ratings over a short period of time, accuracy of results of external criterion measures was improved by removing the chances of outside circumstances non-related to treatment that could alter ratings across time.

An additional strength of the current study was the use of external as compared to self-report criterion measures of psychological symptoms. Past studies have reported inflated effect sizes (e.g., zero-order correlations, ΔR^2 values) due to shared method variance from use of multiple self-report measures (e.g., Sellbom, Graham, et al., 2006; Simms et al., 2005). The current study's use of other-report criterion measures provided more accurate and un-biased estimates of the utility of the RC and Clinical Scales in prediction of psychological symptoms. Furthermore, through the use of three different external criterion measures, the current study was able to lay a foundation of evidence that well estimated of the utility of the RC and Clinical Scales.

Despite these strengths, the study was not without its limitations. Both the Patient Description Form (PDF) and the Spouse Rating Form (SRF) are not commonly used assessments with little evidence from prior studies to attest to their construct validity. Because both of these measures have not been thoroughly tested or validated, generalization of current study results to other research may be limited. For example, the PDF version used in the current study was an abbreviated/modified version of the Patient Description Form (PDF)-Therapist Rating Form developed by Graham, et al. (1999). To create this measure, one researcher selected items to be included on each scale based on similarity of symptoms. With only one reviewer involved, it is possible that item selection might have been more biased based on personal experience than that of the original PDF created by Graham and colleagues in which multiple reviewers were involved in this process. Similarly, only two authors (those of the current study) selected items to be included in SRF Scales based on rational approach, which indicates presence of possible individual bias in item selection. Construct validity of SRF Scales was especially questionable, as internal consistency (Cronbach's alpha) coefficients of the SRF Scales were relatively poor. In conclusion, the results of the current study are limited in ability to be compared to previous studies due to use of unstandardized criterion measures.

Future Directions and Conclusions

In the future, additional research should be conducted using criterion measures that have solid psychometric properties. Use of these additional measures would substantiate evidence showing the possible impact of culture on comparative utility of RC1 with that of Hs. As Hs consistently outperformed RC1 in convergent and discriminant validity, and incremental utility in a majority of external criterion measures (with the exception of convergent and discriminant validity of diagnosis of Somatoform Disorder), there is already substantial evidence showing

impact of cultural differences in somatic reporting on utility of symptom homogeneous versus symptom heterogeneous scales. Further research at the item level could identify symptoms of Hs that were especially beneficial in the Korean culture that might aid clinicians in the future when treating Korean clients.

Due to the inconsistent performance across analyses, RC7 is an additional scale that might benefit from future research using additional criterion measures. Comparing results of convergent and discriminant validity provided no clear patterns of RC7 to Pt. Both RC7 and Pt provided similar convergent and discriminant validity of diagnosis of Anxiety Disorder and discriminant validity of PDF Scales. In contrast, Pt provided greater convergent validity of PDF Scales and greater convergent and discriminant validity of SRF Scales. As suggested by Simms and colleagues (2005), relationships of RC7 with various symptoms of psychological maladjustment are not clear or stable. Research to examine additional correlates is necessary, with the possibility of item content revision to improve assessment utility of symptoms of anxiety. Future studies should also include more diverse diagnostic categories (e.g., Antisocial personality, substance abuse, etc.).

Despite several weaknesses, this study provided substantial evidence for the ability of the RC Scales of the Korean MMPI-2-RF to provide similar to greater utility in assessment of psychological symptoms than that already provided by the Clinical Scales. The relative equivalence between respective RC and Clinical Scale in assessment utility adds support for doubts raised by Bolinskey and Nichols (2011) arguing RC Scales measure different constructs as do their Clinical Scales counterparts, as well as questions regarding their ability to contribute meaningful information to that of respective clinical Scales in prediction of complex and heterogeneous criterion measures (Nichols, 2006). Despite the decrease in item count from

Clinical to RC Scale versions, RC Scales provided similar to slightly improved convergent and discriminant validity, as well as similar to greater proportions of incremental utility across a majority of conceptually relevant criterion measures of psychological symptoms. These findings add support for the comparative efficacy of this newly published and drastically shortened Korean MMPI-2-RF (Han et al., 2011) to its predecessor, the Korean MMPI-2 (Kim et al., 2005). The success of the RC scales to predict a wide variety of psychological symptoms as well as clinical diagnoses in a Korean clinical hospital sample provides evidence that this instrument can be utilized in non-Western (Asian) cultures as substantial similarities were found between study results and previous U.S. MMPI-2-RF research.

The fact that this research was able to usefully predict psychological symptoms in Korean clinical patients should increase the likelihood that clinicians will more frequently use this tool with their clients as an accepted and valued tool of assessing Korean psychological maladjustment. With increased use of the Korean MMPI-2-RF, clinicians should be able to design more effective treatment plans, which will improve the quality of experiences Korean patients have with their therapist. Therefore, Korean clients will more freely express symptoms to their therapists within a professional context. Freedom to express symptoms of psychological distress is extremely difficult within the Korean culture, as past research has argued individuals from Asian cultures are much less likely to seek treatment from professionals as compared to non-Asians due to feelings that admitting to a problem, as well as speaking with a psychologist, will bring shame to themselves or their family (e.g., Braun & Browne, 1998; Kim & Omizo, 2003; Zhang & Dixon, 2003). As more successful treatment outcomes occur due to use of the Korean MMPI-2-RF, perhaps Korean individuals will begin to feel less ashamed for seeking treatment as they realize the benefits of improved quality of life following professional

treatment. As reported by Masuda, Suzumura, Beauchamp, Howells, and Clay (2005), Japanese students who had past experience with seeking psychological professional help had more favorable views of psychological treatment than those without previous treatment. These students were also more likely to understand the need for psychological treatment, as well as to have more confidence in psychological treatment than those who had no past experiences. Applied to Koreans, these positive experiences can lead to increases in (a) the proportion of Korean individuals utilizing psychological treatment services, (b) overall acceptance of psychological treatment within the Korean culture, and (c) improvement in quality of life for many Koreans who previously had not sought treatment due to cultural stigma against therapy.

APPENDICES

APPENDIX A

FREQUENCIES OF CO-MORBIDITY BY PRIMARY DIAGNOSIS OF CLINICAL SAMPLE

Diagnostic Group	Co-morbidity Status	f (%)
MDD	Pure	67 (79%)
	Comorbid	18 (21%)
	Total	85 (100%)
Somatoform	Pure	18 (28%)
	Comorbid	46 (72%)
	Total	64 (100%)
Schizophrenia	Pure	69 (97%)
	Comorbid	2 (3%)
	Total	71 (100%)
Anxiety Disorder	Pure	62 (72%)
	Comorbid	24 (28%)
	Total	86 (100%)
Bipolar Disorder	Pure	75 (84%)
	Comorbid	14 (16%)
	Total	89 (100%)
Total	Pure	291 (74%)
	Comorbid	104 (26%)
	Total	395 (100)%

APPENDIX B

DATA SCREENING AND ASSUMPTION CHECKS

Missing values. Missing values were noted in items of the RC and Clinical Scales. All

missing values were treated as non-keyed responses. For the clinical sample, with the exception

of RC6, which had only one item with one missing value ($n_{missing} = 1$), most of the RC Scales had

five to six items with one missing value. RCd, RC6, and RC7 had one item with three missing

values ($n_{missing} = 3$), while RC1 and RC9 each had one item with two missing values ($n_{missing} = 2$).

RC7 had one item with four missing values ($n_{missing} = 4$). Most Clinical Scales also had several

items with one missing value ($n_{missing} = 1$). Hs, D, and Pt had one item with two missing values

($n_{missing} = 2$). Pa and Sc had two items, while Ma had three items with two missing values ($n_{missing}$

$= 2$). D also had one item with three missing values ($n_{missing} = 3$), and Sc had three items with

three missing values. PDF Scales Depressed, Suspicious, Psychotic Symptoms, and Anxious

contained one item with one missing value ($n_{missing} = 1$). Cases with missing values were not used

in hierarchical regression analyses predicting these four PDF Scales. As a result, these regression

analyses utilized 394 instead of 395 cases. Items of Clinical and RC Scales contained no missing

values ($n_{missing} = 0$) for the Korean normative couple sample.

Equivalence of groups. One-way ANOVAs and chi-square tests for independence were

used to determine whether the five diagnostic groups were equivalent in terms of demographic

variables for the clinical sample. Results indicated that the five groups did differ significantly in

terms of mean age, $F(4, 390) = 14.83$, $p < .001$ and mean years of education attained, $F(4, 390)$

$= 3.17$, $p < .05$. Using Tukey's HSD Test, patients diagnosed with Schizophrenia were on

average significantly younger than those diagnosed with Depressive, Anxiety, or Somatoform

Disorders. Patients diagnosed with Bipolar Disorder were also younger than those diagnosed

with Depressive, Anxiety, or Somatoform Disorders, on average. Patients diagnosed with Bipolar Disorder had received significantly more years of education, on average, than those with Somatoform Disorder.

Chi-square analyses revealed significant relationship between diagnostic status and gender χ^2 (4, $n = 395$) = 13.68, $p < .01$. There were more females than males diagnosed with Depressive, Bipolar, and Somatoform Disorders, while there were more males than females diagnosed with Anxiety Disorder. There was also a significant relationship between diagnostic status and marital status, χ^2 (20, $n = 395$) = 64.12, $p < .001$. For all diagnostic categories, the majority of patients reported being never married or married. Very few reported marital status of divorced, separated, widowed, or remarriage. For individuals diagnosed with Schizophrenia and Bipolar Disorder, most were never married, whereas most individuals diagnosed with Depressive, Anxiety, or Somatoform Disorders reported being married. There was also a significant relationship between occupational and diagnostic status, χ^2 (32, $n = 395$) = 104.02, $p < .001$. Individuals diagnosed with Schizophrenia and Bipolar Disorders most frequently reported being a student, while individuals diagnosed with Depressive, Anxiety, and Somatoform Disorders most frequently reported occupation as housewife.

In the Korean normative couple sample, results of one-way ANOVAs by gender indicated significant mean differences in age, F (1, 214) = 4.85, $p < .05$ and years of education, F (1, 214) = 4.91, $p < .05$. Males were significantly older and had more years of education than females. Results of chi-square analyses showed no significant relationship between gender and religion, χ^2 (4, $n = 215$) = 7.43, $p > .05$.

Assumption Checks

Outliers: Logistic regression. In the clinical sample, Cook's measure of influence indicated in Cook's *D* values showed that no cases had *D* values > 1.0. Cook's distance values for all models were less than .15, with the exception of one case in the models predicting Bipolar Disorder from 1) Ma and RC9 and 2) Ma, RC9, and RCd, in which Cook's *D* value was .16 in both models.

Several logistic regression standardized residuals were deemed outliers within each model predicting diagnostic status. Of the logistic regression models performed predicting diagnostic status versus all other diagnoses, one case surpassed the $z \geq 3.29$ criterion in prediction of Major Depressive Disorder in the models using (a) D and RCd and (b) D and RC2 as predictors. One case also surpassed the criterion in prediction of Somatoform Disorder from Hs and RC1, while two cases exceeded the criterion in prediction of diagnosis of Schizophrenia from Sc and RC8. No cases predicting diagnosis of Anxiety Disorder from Pt and RC7 or predicting diagnosis of Schizophrenia from Pa and RC6 exceeded criterion. In prediction of diagnosis of Bipolar Disorder from Ma and RC9, six cases exceeded the criterion.

In the models predicting diagnostic status in which RCd was added as a predictor in the RC Scale step, four cases exceeded the $z \geq 3.29$ criterion in prediction of diagnosis of Major Depressive Disorder from RC2, RCd, and D. In the model predicting diagnosis of Somatoform Disorder from RC1, RCd, and Hs, two cases had z-scores exceeding the criterion, while five cases exceeded the criterion in the model predicting diagnosis of Schizophrenia from Pa, RC6, and RCd. Four cases exceeded the criterion in the model predicting diagnosis of Schizophrenia from Sc, RC8, and RCd. Only one case exceeded the criterion in the model predicting diagnosis of Anxiety Disorder from Pt, RC7, and RCd. Five cases had z-scores above the criterion in the

83

model predicting diagnosis of Bipolar Disorder from Ma, RC9, and RCd. Due to the large sample size ($N = 395$), a few exceptional cases exceeding the $z \geq 3.29$ criterion were expected; therefore, no cases were eliminated.

Outliers: Multiple regression. In the clinical sample, standardized residuals from models predicting PDF scale scores were also examined for extremely large scores with $z \geq 3.29$ criterion (Tabachnick & Fidell, 2007). Nine cases were observed with z-scores ranging between 3.30 and 3.80. One case in the model predicting Suspicious from Pa and RC6, two cases in the model predicting Psychotic from Sc and RC8, three cases in the model predicting Narcissistic from Ma and RC9, and three cases in the model predicting Agitated from Ma and RC9. An additional five cases with extreme z-scores were found in the models with RCd entered simultaneously with the RC Scale in prediction of PDF Scales. Models included one case in the model predicting Suspicious from Pa, RC6, and RCd, one case in the model predicting Psychotic from Sc, RC8, and RCd, and three cases in the model predicting Agitated from Ma, RC9, and RCd.

Standardized residuals from models predicting SRF Scale scores in the Korean normative couple sample were also examined for outliers. The following models had cases with z-scores exceeding the cut-off criterion ($\geq \pm 3.29$): one case in the model predicting Hs_R from Hs and RC1, three cases in the model prediction D_R from D and RCd, two cases in the model predicting D_R from D and RC2, and one case in the model predicting Pa_R from Pa and RC6.

In the models with RCd entered simultaneously with the RC Scale in prediction of SRF Scales, the following had cases with z-scores exceeding the criterion: one case in the model predicting Hs_R from RC1, RCd, and Hs, three cases in the model predicting D_R from RCd, RC2, and D, two cases in the model predicting Pa_R from RC6, RCd, and Pa, and one case in the

model predicting Pt_R from RCd, RC7, and Pt. With a large example, a small number of extreme cases are expected (Tabachnick & Fidell, 2007); therefore, these cases were not removed.

Multicollinearity: Logistic regression. Model parameters and standard errors for each of the hierarchical logistic regression models (with and without RCd) were examined for extremely large values indicating presence of multicollinearity. No multicollinearity was assumed due to there being no extremely large parameter or standard error values.

Multicollinearity and singularity: Multiple regression. The assumption of singularity (tolerance values greater than 0) and multicollinearity (condition indices from collinearity diagnostics < 30 and not more than two variables with variance proportions >.50) was met for all models in the clinical sample. Similarly, no models failed this assumption in models in which RCd was entered simultaneously with the RC Scale in prediction of the conceptually relevant PDF scale. The assumption of singularity and multicollinearity using the same criteria as in the clinical sample was met for all models predicting SRF Scales in the Korean normative couple sample. No models had condition index values >30. Similarly, no models failed this assumption in models in which RCd was entered simultaneously with the RC Scale in prediction of the conceptually relevant SRF scale.

Linearity: Logistic regression. Of the 14 hierarchical logistic regression models performed in the clinical sample in which a Clinical and RC Scale was entered in step one, and an interaction term between each RC or Clinical Scale and its natural logarithm was entered in step two, no interaction terms were significant. Of the 12 models predicting diagnostic category in which RCd was added to the model, the interaction term of RCd and its natural logarithm was significant in the model predicting diagnosis of Schizophrenia from Sc, RC8, and RCd ($p = .05$). All other models met the assumption and demonstrated a linear relationship with the logit

transformation of the dependent variable and the predictor variables. As only one model failed the assumption, no variables were transformed to ease interpretation.

Normality, linearity, homoscedasticity: Multiple regression. Normality, linearity, and homoscedasticity assumptions were examined through scatter-plots of residuals against predicted dependent variable (e.g., PDF) scores in the clinical sample. These assumptions were met for all models with the exception of the model predicting Psychotic Symptoms from RC8 and Sc. Analyses of scatter plots of residuals by predicted values yielded non-random scatter, and a frequency histogram of residuals showed positive skew. In the models in which RCd was added in the RC Scale step, the model predicting Psychotic Symptoms from Sc, RC8, and RCd, yielded non-random scatter plot, a severe positive skew frequency histogram of residuals, and extreme curvilinear trend Normal P-P plot of observed by expected cumulative probabilities.

Failure of assumptions of normality and homoscedasticity was likely due to the fact that RC8 was significantly and positively skewed. Transformations (e.g., square root, logarithmic) are commonly performed to correct for failures of these assumptions. Transformation of variables makes interpretation of results more difficult, and is not always recommended (Tabachnick & Fidell, 2007). Transformations that would encumber interpretation of results were not performed on RC8.

Normality, linearity, and homoscedasticity assumptions were examined through scatter-plots of residuals against predicted dependent variable (e.g., SRF) scores in the Korean normative couple sample as well. Non-random scatter was noted in the model predicting Sc_R from Sc and RC8. In the models in which RCd was entered simultaneously with the RC scale, the model predicting Pa_R from RC6, RCd, and Pa, showed slight non-random scatter, and a slight positive skew histogram. Most other scatter plots showed slight deviation to random

86

scatter with a few outliers. Transformations were not performed due to the relatively few models failing this assumption.

Independence of error. Multilevel modeling was conducted to examine the impact of therapist variables on client ratings in the clinical sample. This method determines the degree to which the assumption of independence of errors has been violated by providing an assessment the degree of relatedness (intraclass correlation coefficient; ICC). Variances between groups are compared to the variances within groups with large ICC values indicating greater violation of independence of errors and greater chances of Type I error if adjustments are not made for level of relatedness (Tabachnick & Fidell, 2008). For example, Barcikowski (1981) estimate with an alpha of .05 and ICC of .05, the actual significance level is .11. Unfortunately, there are no strict rules of thumb for how large ICC should be in order to deem MLM necessary. Adelson and Owen (2012) suggest using MLM when ICC is as small as .05 indicating 5% of rating score variability is a result of differences between therapists. Tabachnick and Fidell (2008) suggest running the analysis both ways. If results do not differ substantially, report findings using the simplest method possible.

Past research has reported difficulties (e.g., non-converging models) with use of multi-level modeling due to the varied proportions of SCID diagnostic categories commonly observed among different therapists (Wampold & Brown, 2005). Proportions of SCID diagnoses also varied substantially among therapists of the current study: out of fifteen therapists, three had no clients diagnosed with Schizophrenia, and four therapists had no clients diagnosed with Somatoform Disorder. In comparison, 50% of clients assigned to two therapists represented diagnoses of Schizophrenia or Somatoform Disorder, respectively. MLM was therefore performed using only PDF Scales as the criterion measure.

87

To determine if MLM is required, the first step is to run the intercepts only model (without predictors). This model assesses group differences in means (intercepts). For example, scores on each dependent variable (conceptually relevant PDF Scale) were predicted by an intercept that was allowed to vary across groups with therapists being considered a random effect (Tabachnick & Fidell, 2008). The resulting model provides an estimate of the ICC ranging from 0 to 1, or the estimate of similarity of PDF ratings of clients from the same therapist (Adelson & Owen 2012). This model also tests the differences in intercepts (means) using a z-score. With a one-tailed test, all models with intercept z-scores > 1.58 indicate that group differences should be taken into account. Tabachnick and Fidell suggest using these z-score tests, along with ICC values, to evaluate need for ICC (2008). ICC values for PDF Scales were as follows: Somatic (.06), Depressed (.15), Suspicious (.05), Anxious (.07), Obsessive-Compulsive (.14), Insecure (.08), Psychotic (.12), Agitated (.13), and Narcissistic (.14). In addition, the following models yielded z-scores > 1.58: Obsessive-compulsive (z-score = 1.91), Psychotic (z-score = 1.79), Agitated (z-score = 1.66), and Narcissistic (z-score = 1.94).

As all models yielded ICC values ≥ .05, the next step of MLM was performed for all conceptually relevant PDF Scales. In this step, a predictor was added to the model to examine if the effects of the predictor (e.g., RC or Clinical Scale) on conceptually relevant PDF Scale rating differed significantly depending on groups (therapists). Only the model predicting PDF Scale Anxious from Pt indicated a significant difference depending on therapist (z-score = -2.98, $p <$ .01). Separate linear regression models predicting Anxious from Pt were then performed for each therapist. In 13 of the 15 models (therapists), Anxious and Pt showed a moderate positive correlation ($r_{mean} = .37$, $r_{median} = .35$). In contrast, results of zero-order correlations of Anxious with Pt from two therapists indicated strong negative relationships ($rs = -.99$ and -.65). Since

these two therapists had relatively few clients (3 and 7 clients, respectively), resulting correlation coefficients were most likely unstable. The impact of violation in independence of errors was considered negligible with PDF Scale ratings assumed to be minimally impacted by differences due to therapists. Results were reported without adjustment, as suggested by Tabachnick and Fidell (2008).

APPENDIX C

SUMMARY OF DISCRIMINANT VALIDITY CALCULATIONS
STRUCTURED CLINICAL INTERVIEW FOR DSM-IV (SCID)

Scale	Conceptually Relevant Criteria (r)	Conceptually Unrelated Criteria r_{mean} (absolute value r)	Discriminant validity (r_{dis})
RCd	MDD $r = .15$	([SOM $r = .08$] + [SCZ $r = .09$] +[ANX $r = .13$] + [BP $r = .13$])/4 =.11	RCd $r_{dis} = .15-.11 = .04$
D	MDD $r = .22$	([SOM $r = .13$] + [SCZ $r = .15$] +[ANX $r = .01$] + [BP $r = .21$])/4 = .13	D $r_{dis} = .22-.13 = .09$
RC1	SOM $r = .17$	([MDD $r = .02$] + [SCZ $r = .18$] +[ANX $r = .09$] + [BP $r = .08$])/4 =.09	RC1 $r_{dis} = .17-.09 = .08$
Hs	SOM $r = .19$	([MDD $r = .08$] + [SCZ $r = .22$] +[ANX $r = .06$] + [BP $r = .02$])/4 =.10	Hs $r_{dis} = .19-.10 = .09$
RC2	MDD $r = .17$	([SOM $r = .08$] + [SCZ $r = .05$] +[ANX $r = .01$] + [BP $r = .21$])/4 =.09	RC2 $r_{dis} = .17-.09 = .08$
D	MDD $r = .22$	([SOM $r = .13$] + [SCZ $r = .15$] +[ANX $r = .01$] + [BP $r = .21$])/4 = .13	D $r_{dis} = .22-.13 = .09$
RC6	SCZ $r = .16$	([MDD $r = .04$] + [SOM $r = .13$] +[ANX $r = .13$] + [BP $r = .13$])/4 =.11	RC6 $r_{dis} = .16-.11 = .05$
Pa	SCZ $r = .12$	([MDD $r = .00$] + [SOM $r = .10$] +[ANX $r = .18$] + [BP $r = .15$])/4 =.11	Pa $r_{dis} = .12-.11 = .01$
RC7	ANX $r = -.03$	([MDD $r = .04$] + [SOM $r = .13$] +[SCZ $r = .10$] + [BP $r = .20$])/4 =.12	RC7 $r_{dis} = -.03-.12 = -.15$
Pt	ANX $r = -.04$	([MDD $r = .11$] + [SOM $r = .07$] +[SCZ $r = .13$] + [BP $r = .11$])/4 =.11	Pt $r_{dis} = -.04-.11 =-.15$
RC8	SCZ $r = .11$	([MDD $r = .06$] + [SOM $r = .13$] +[ANX $r = .15$] + [BP $r = .21$])/4 =.14	RC8 $r_{dis} = .11-.14 = -.03$
Sc	SCZ $r = -.01$	([MDD $r = .05$] + [SOM $r = .12$] +[ANX $r = .15$] + [BP $r = .21$])/4 =.13	Sc $r_{dis} = -.01-.13 = -.14$
RC9	BP $r = .36$	([MDD $r = .07$] + [SOM $r = .17$] +[SCZ $r = .06$] + [ANX $r = .08$])/4 =.10	RC9 $r_{dis} = .36-.10 = .26$
Ma	BP $r = .29$	([MDD $r = .02$] + [SOM $r = .13$] +[SCZ $r = .02$] + [ANX $r = .14$])/4 =.08	Ma $r_{dis} = .29-.08 = .21$

APPENDIX D

SUMMARY OF DISCRIMINANT VALIDITY CALCULATIONS
PATIENT DESCRIPTION FORM (PDF)

Scale	Relevant Criteria (r)	Conceptually Unrelated Criteria r_{mean} (absolute value r)	Discriminant validity (r_{dis})
RCd	Dep r = .56	([Som r = .12] + [Susp r = .09] +[Anx r = .22] + [Insec r = .41] + [Obs r =.15] +[Psych r =.06] +[Narc r = .06] + [Agit r = .04])/8 =.14	RCd r_{dis} = .56-.14 = .42
D	Dep r = .52	([Som r = .39] + [Susp r = .15] +[Anx r = .31] + [Insec r = .29] + [Obs r =.20] +[Psych r =.25] +[Narc r = .17] + [Agit r = .05])/8 =.23	D r_{dis} = .52-.23 = .29
RC1	Som r = .39	([Dep r = .38] + [Susp r = .01] +[Anx r = .15] + [Insec r = .14] + [Obs r =.04] +[Psych r =.11] +[Narc r = .02] + [Agit r = .10])/8 =.12	RC1 r_{dis} = .39-.12 = .27
Hs	Som r = .48	([Dep r = .44] + [Susp r = .05] +[Anx r = .20] + [Insec r = .15] + [Obs r =.00] +[Psych r =.19] +[Narc r = .01] + [Agit r = .09])/8 =.14	Hs r_{dis} = .48-.14 = .34
RC2	Dep r = .46	([Som r = .22] + [Susp r = .10] +[Anx r = .25] + [Insec r = .35] + [Obs r =.19] +[Psych r =.16] +[Narc r = .25] + [Agit r = .13])/8 =.21	RC2 r_{dis} = .46-.21 = .25
D	Dep r = .52	([Som r = .39] + [Susp r = .15] +[Anx r = .31] + [Insec r = .29] + [Obs r =.20] +[Psych r =.25] +[Narc r = .17] + [Agit r = .05])/8 =.23	D r_{dis} = .52-.23 = .29
RC6	Susp r = .34	([Dep r = .20] + [Som r = .09] +[Anx r = .04] + [Insec r = .19] + [Obs r =.01] +[Psych r =.34] +[Narc r = .06] + [Agit r = .19])/8 =.14	RC6 r_{dis} = .34-.14 = .20
Pa	Susp r = .33	([Dep r = .39] + [Som r = .00] +[Anx r = .09] + [Insec r = .30] + [Obs r =.09] +[Psych r =.27] +[Narc r = .06] + [Agit r = .19])/8 =.17	Pa r_{dis} = .33-.17 = .16
RC7	Anx r = .20	([Dep r = .38] + [Som r = .03] +[Susp r = .11] + [Insec r = .28] + [Obs r =.12] +[Psych r =.02] +[Narc r = .02] + [Agit r = .11])/8 =.13	RC7 r_{dis} = .20-.13 = .07
Pt	Anx r = .26	([Dep r = .51] + [Som r = .17] +[Susp r = .07] + [Insec r = .35] + [Obs r =.08] +[Psych r =.09] +[Narc r = .07] + [Agit r = .07])/8 =.17	Pt r_{dis} = .26-.18 = .08
RC7	Insec r = .28	([Dep r = .38] + [Som r = .03] +[Susp r = .11] + [Anx r = .20] + [Obs r =.12] +[Psych r =.02] +[Narc r = .02] + [Agit r = .11])/8 =.12	RC7 r_{dis} = .28-.12=.16
Pt	Insec r = .35	([Dep r = .51] + [Som r = .17] +[Susp r = .07] + [Anx r = .26] + [Obs r =.08] +[Psych r =.09] +[Narc r = .07] + [Agit r = .07])/8 =.17	Pt r_{dis} = .35-.17 = .18
RC7	Obs r = .08	([Dep r = .38] + [Som r = .03] +[Susp r = .11] + [Anx r = .20] + [Insec r =.28] +[Psych r =.02] +[Narc r = .02] + [Agit r = .11])/8 =.14	RC7 r_{dis} = .08-.14 = -.06
Pt	Obs r = .12	([Dep r = .51] + [Som r = .17] +[Susp r = .07] + [Anx r = .26] + [Insec r =.35] +[Psych r =.09] +[Narc r = .07] + [Agit r = .07])/8 =.20	Pt r_{dis} = .12-.20 = -08
RC8	Psych r = .28	([Dep r = .26] + [Som r = .02] +[Susp r = .29] + [Anx r = .04] + [Insec r =.15] +[Obs r =.01] +[Narc r = .11] + [Agit r = .19])/8 =.13	RC8 r_{dis} = .28-.13 = .15
Sc	Psych r = .09	([Dep r = .49] + [Som r = .07] +[Susp r = .22] + [Anx r = .15] + [Insec r =.36] +[Obs r =.06] +[Narc r = .04] + [Agit r = .14])/8 =.19	Sc r_{dis} = .09-.19 = -.10
RC9	Narc r = .19	([Dep r = .04] + [Som r = .13] +[Susp r = .19] + [Anx r = .04] + [Insec r =.05 +[Obs r =.10] +[Psych r = .06] + [Agit r = .21])/8 =.10	RC9 r_{dis} = .19-.10 = .09
Ma	Narc r = .19	([Dep r = .08] + [Som r = .11] +[Susp r = .19] + [Anx r = .08] + [Insec r =.02] +[Obs r =.11] +[Psych r = .14] + [Agit r = .17])/8 =.11	Ma r_{dis} = .19-.11 = .08
RC9	Agit r = .21	([Dep r = .04] + [Som r = .13] +[Susp r = .19] + [Anx r = .04] + [Insec r =.05 +[Obs r =.10] +[Psych r = .06] + [Narc r = .19])/8 =.10	RC9 r_{dis} = .21-.10 = .11
Ma	Agit r = .17	([Dep r = .08] + [Som r = .11] +[Susp r = .19] + [Anx r = .08] + [Insec r =.02] +[Obs r =.11] +[Psych r = .14] + [Narc r = .19])/8 =.12	Ma r_{dis} = .17-.12 = .05

APPENDIX E

SUMMARY OF DISCRIMINANT VALIDITY CALCULATIONS
PATIENT DESCRIPTION FORM (PDF)

Scale	Conceptually Relevant Criteria (r)	Conceptually Unrelated Criteria r_{mean} (absolute value r)	Discriminant validity (r_{dis})
RCd	D_R r = .37	([Hs_R r = .39] + [Pa_R r = .31] +[Pt_R r = .32] + [Sc_R r = .04] + [Ma_R r = .16])/5=.24	RCd r_{dis} = .37-.24 = .13
D	D_R r = .19	([Hs_R r = .30] + [Pa_R r = .08] +[Pt_R r = .21] + [Sc_R r = .10] + [Ma_R r .06])/5=.11	D r_{dis} = .19-.15 = .04
RC1	Hs_R r = .25	([D_R r = .15] + [Pa_R r = .16] +[Pt_R r = .16] + [Sc_R r = .03] + [Ma_R r = .08])/5=.12	RC1 r_{dis} = .25-.12 = .13
Hs	Hs_R r = .34	([D_R r = .21] + [Pa_R r = .18] +[Pt_R r = .20] + [Sc_R r = .04] + [Ma_R r = .07])/5=.14	Hs r_{dis} = .34-.14 = .20
RC2	D_R r = .09	([Hs_R r = .21] + [Pa_R r = .03] +[Pt_R r = .12] + [Sc_R r = .08] + [Ma_R r = .18])/5=.12	RC2 r_{dis} = .09-.12 =-.03
D	D_R r = .19	([Hs_R r = .30] + [Pa_R r = .08] +[Pt_R r = .21] + [Sc_R r = .10] + [Ma_R r .06])/5=.11	D r_{dis} = .19-.15 = .04
RC6	Pa_R r = .08	([D_R r = .03] + [Hs_R r = .03] +[Pt_R r = .06] + [Sc_R r = .14] + [Ma_R r = .12])/5=.13	RC6 r_{dis} = .08-.08 = .00
Pa	Pa_R r = .19	([D_R r = .16] + [Hs_R r = .17] +[Pt_R r = .17] + [Sc_R r = .01] + [Ma_R r = .17])/5=.14	Pa r_{dis} = .19-.14 = .05
RC7	Pt_R r = .15	([D_R r = .22] + [Hs_R r = .25] +[Pa_R r = .23] + [Sc_R r = .02] + [Ma_R r = .16])/5=.18	RC7 r_{dis} = .15-.18 =-.03
Pt	Pt_R r = .27	([D_R r = .29] + [Hs_R r = .34] +[Pa_R r = .25] + [Sc_R r = .03] + [Ma_R r = .13])/5=.21	Pt r_{dis} = .27-.21 = .06
RC8	Sc_R r = .05	([D_R r = .17] + [Hs_R r = .13] +[Pa_R r = .19] + [Pt_R r = .09] + [Ma_R r = .21])/5=.16	RC8 r_{dis} = .05-.16 = -.11
Sc	Sc_R r = -.02	([D_R r = .25] + [Hs_R r = .30] +[Pa_R r = .24] + [Pt_R r = .19] + [Ma_R r = .18])/5=.23	Sc r_{dis} = -.02-.23 = -.25
RC9	Ma_R r = .23	([D_R r = .11] + [Hs_R r = .09] +[Pa_R r = .13] + [Pt_R r = .04] + [Sc_R r = .04])/5=.08	RC9 r_{dis} = .23-.08 = .15
Ma	Ma_R r = .17	([D_R r = .09] + [Hs_R r = .05] +[Pa_R r = .14] + [Pt_R r = .10] + [Sc_R r = .00])/5=.08	Ma r_{dis} = .17-.08 = .09

APPENDIX F

LOGISTIC REGRESSION ANALYSES RESULTS FOR PREDICTING DIAGNOSIS
STATUS FROM CONCEPTUALLY RELEVANT RESTRUCTURED CLINICAL (RC),
CLINICAL SCALES, AND RCD

Step	Scale	β	OR	Model fit χ^2	$\Delta\chi^2$	R^2	ΔR^2	$\Delta\chi^2_1$	ΔR^2_1
MDD									
Step 1	D	$.08^*$	1.08	21.22^{***}		.081			
Step 2	RC2	.01	1.01						
	RCd	.02	1.02	21.76^{***}	.54	.083	.002	6.72^*	.025
SOM									
Step 1	Hs	$.20^{**}$	1.22	14.68^{***}		.062			
Step 2	RC1	-.04	.96						
	RCd	$-.13^{***}$	.88	36.74^{***}	22.06^{***}	.151	.089	8.02^{**}	.032
SCZ									
Step 1	Pa	.08	1.09	5.94^*		.024			
Step 2	RC6	.12	1.12						
	RCd	$-.12^{***}$	.89	26.97^{***}	21.04^{***}	.108	.084	2.90	.011
ANX									
Step 1	Pt	$.08^*$	1.08	.73		.003			
Step 2	RC7	.03	1.03						
	RCd	$-.18^{***}$	.83	17.05^{**}	16.32^{***}	.065	.062	5.11^*	.019
SCZ									
Step 1	Sc	-.02	.98	.06		.000			
Step 2	RC8	$.18^{**}$	1.20						
	RCd	-.07	.94	18.08^{***}	18.02^{***}	.073	.073	.70	.002
BP									
Step 1	Ma	.01	1.01	33.28^{***}		.123			
Step 2	RC9	$.16^{***}$	1.17						
	RCd	.00	1.00	52.44^{***}	19.16^{***}	.190	.067	.02	.001

Note. β = standardized regression coefficients in the full model where all predictors are entered. OR=Odds Ratio (e^b). Nagelkerke R^2 was used. $\Delta\chi^2_1$ = Incremental χ^2 added by Clinical Scale over and above RCX. ΔR^2_1 = Incremental R^2 added by Clinical Scale over and above RCX. MDD = Major Depressive Disorder; SOM = Somatoform Disorder; SCZ = Schizophrenia; ANX = Anxiety Disorder; BP = Bipolar Disorder; RCd = Demoralization; RC1 = Somatic Complaints; RC2 = Low Positive Emotions; RC6 = Ideas of Persecution; RC7 = Dysfunctional Negative Emotions; RC8 = Aberrant Experiences; RC9 = Hypomanic Activation; Hs = Hypochondriasis; D = Depression; Pa = Paranoia; Pt = Psychasthenia; Sc = Schizophrenia; Ma = Hypomania.
$^* = p < .05.$ $^{**} = p < .01.$ $^{***} = p < .001.$

PREDICTION OF PATIENT DESCRIPTION FORM SCALES FROM RCD AND CONCEPTUALLY RELEVANT RC AND CLINICAL SCALES

Step	Scale	R^2	ΔR^2	β	r	sr^2_1
Depressed						
Step 1	D	.271***		.21**	.52***	
Step 2	RC2			.13*	.46***	
	RCd	.389***	.118***	.39***	.56***	.017**
Somatic						
Step 1	Hs	.231***		1.04***	.48***	
Step 2	RC1			-.48***	.39***	
	RCd	.289***	.058***	-.20***	.12*	.122***
Suspicious						
Step 1	Pa	.106***		.26**	.33***	
Step 2	RC6			.21**	.34***	
	RCd	.141***	.035***	-.17**	.09	.023**
Anxious						
Step 1	Pt	.066***		.37**	.26***	
Step 2	RC7			-.09	.20***	
	RCd	.068***	.002	-.04	.22***	.018**
Insecure						
Step 1	Pt	.126***		.05	.35***	
Step 2	RC7			-.13	.28***	
	RCd	.173***	.047***	.47***	.41***	.000
Obs-Comp						
Step 1	Pt	.015*		.05	.12*	
Step 2	RC7			-.12	.08	
	RCd	.027*	.012	.20	.15**	.001
Psychotic						
Step 1	Sc	.008		-.07	.09	
Step 2	RC8			.45***	.28***	
	RCd	.132***	.124***	-.24**	-.06	.001
Narcissistic						
Step 1	Ma	.037***		.13	.19***	
Step 2	RC9			.15	.19***	
	RCd	.063***	.025**	-.16**	-.06	.007
Agitated						
Step 1	Ma	.029**		.01	.17**	
Step 2	RC9			.22**	.21***	
	RCd	.047***	.018*	-.04	.04	.000

Note. β = standardized regression coefficients in the full model where all predictors are entered. r = zero-order correlation of each predictor with a criterion. sr^2_1 = Incremental R^2 added by Clinical Scale X over and above RCX (and RCd). RC = Restructured Clinical Scale. RCd = Demoralization; RC1 = Somatic Complaints; RC2 = Low Positive Emotions; RC6 = Ideas of Persecution; RC7 = Dysfunctional Negative Emotions; RC8 = Aberrant Experiences; RC9 = Hypomanic Activation; Hs = Hypochondriasis; D = Depression; Pa = Paranoia; Pt = Psychasthenia; Sc = Schizophrenia; Ma = Hypomania; Obs-Comp = Obsessive-Compulsive

PREDICTION OF SPOUSE RATING FORM (SRF) SCALES FROM RCD AND CONCEPTUALLY RELEVANT RC AND CLINICAL SCALES

Step	Scale	R^2	ΔR^2	β	r	sr^2_1
D_R						
Step 1	D	.035**		.03	.19**	
Step 2	RC2			-.01	.09	
	RCd	.139***	.104***	.36***	.37***	.000
Hs_R						
Step 1	Hs	.113***		.68***	.34***	
Step 2	RC1			-.54**	.25***	
	RCd	.215***	.102***	.31***	.39***	.059***
Pa_R						
Step 1	Pa	.034**		.14	.19**	
Step 2	RC6			-.15	.08	
	RCd	.107***	.073***	.30***	.31***	.009
Pt_R						
Step 1	Pt	.074***		.17	.27***	
Step 2	RC7			-.28*	.15*	
	RCd	.125***	.051**	.39**	.32***	.005
Sc_R						
Step 1	Sc	.001		-.33*	-.02	
Step 2	RC8			.20	.05	
	RCd	.024	.023	.18	.04	.021*
Ma_R						
Step 1	Ma	.030*		.04	.17*	
Step 2	RC9			.18	.23**	
	RCd	.057**	.027	.06	.16*	.001

Note. β = standardized regression coefficients in the full model where all predictors are entered. r = zero-order correlation of each predictor with a criterion. sr^2_1 = Incremental R^2 added by Clinical Scale X over and above RCX (and RCd). RC = Restructured Clinical Scale. RCd = Demoralization; RC1 = Somatic Complaints; RC2 = Low Positive Emotions; RC6 = Ideas of Persecution; RC7 = Dysfunctional Negative Emotions; RC8 = Aberrant Experiences; RC9 = Hypomanic Activation; Hs = Hypochondriasis; D = Depression; Pa = Paranoia; Pt = Psychasthenia; Sc = Schizophrenia; Ma = Hypomania

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

REFERENCES

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

Barcikowski, R. S. (1981). Statistical power with group mean as the unit of analysis. *Journal of Educational Statistics, 6*(3), 267-285. doi: 10.2307/1164877

Barthlow, D. L., Graham, J. R., Ben-Porath, Y. S., & McNulty, J. L. (1999). Incremental validity of the MMPI-2 Content Scales in an outpatient mental health setting. *Psychological Assessment, 11*(1), 39-47.

Ben-Porath Y. S. (2012). *Interpreting the MMPI-2-RF*. Minneapolis, MN: University of Minnesota Press.

Ben-Porath, Y. S., & Tellegen, A. (2008). *MMPI-2-RF (Minnesota Multiphasic Personality Inventory—2 Restructured Form): Manual for administration, scoring, and interpretation.* Minneapolis, MN: University of Minnesota Press.

Bolinskey, P. K. & Nichols, D. S. (2011). Construct drift in the MMPI-2 Restructured Clinical Scales: Further evidence and a possible historic example. *Journal of Clinical Psychology, 67*(9), 907-917.

Butcher, J. N., Graham, J. R., & Ben-Porath, Y. S. (1995). Methodological problems and issues in MMPI, MMPI–2, and MMPI–A research. *Psychological Assessment*, *7*, 320–329.

Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W. G., & Kaemmer, B. (1989). *MMPI-2 (Minnesota Multiphasic Personality Inventory-2) manual for administration, scoring, and interpretation* (Rev. ed.). Minneapolis, MN: University of Minnesota Press.

Butcher, J. N., Hamilton, C. K., Rouse, S. V., & Cumella, E. J. (2006). The deconstruction of the Hy Scale of the MMPI-2:Failure of RC3 in measuring somatic symptom expression. *Journal of Personality Assessment, 87*(2), 186-192.

Butcher, J.N., Lim, J., & Nezami, E. (1998). Objective study of abnormal personality in cross-cultural settings: The Minnesota Multiphasic Personality Inventory (MMPI-2). *Journal of Cross-Cultural Psychology. Special Issue: Personality and its measurement in cross-cultural perspective, 29*(1), 189-211.

Braun, K. S. & Browne, C. V. (1998). Perceptions of dementia, caregiving, and help seeking among Asian and Pacific Islander American. *Health & Social Work, 23* (4), 262-274.

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

Cohen, Jacob (1988). *Statistical Power Analysis for the Behavioral Sciences* (second ed.). New York City, New York: Lawrence Erlbaum Associates.

Derogatis, L. R., Rickels, K., & Rock, A. (1976). The SCL–90 and the MMPI: A step in the validation of a new self-report scale. *British Journal of Psychiatry, 128*, 280–289.

Doverspike, W. F. (1990). *Multiaxial Diagnostic Inventory*. Atlanta, GA: Author.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1997). *Structured clinical interview for Axis I DSM-IV disorders* (SCID). New York: New York State Psychiatric Institute Biometric Research Department.

Forbey, J. D., Lee, T. T. C., & Handel, R. W. (2010). Correlates of the MMPI-2-RF in a college setting. *Psychological Assessment, 22*(4), 737-744.

Graham, J. R. (2006). *MMPI-2: Assessing personality and psychopathology* (4th ed.). Oxford, NY: Oxford University Press.

Graham, J. R., Ben-Porath, Y. S., & McNulty, J. L. (1999). *MMPI–2 correlates for outpatient mental health settings.* Minneapolis: University of Minnesota Press.

Han, K. (1993). The use of the MMPI-2 in Korea: Inventory adaptation, equivalence evaluation, and initial validation. *Dissertation Abstracts International, 54* (5-B).

Han, K. (1996). Adaptations in Asia and Southeast Asia. In J. N. Butcher (ed.), *International adaptations of the MMPI-2: Research and clinical applications* (67-218). Minneapolis, MN: University of Minnesota Press.

Han, K., Moon, K., Lee, J., & Kim, J. (2011). *Korean MMPI-2-RF user manual.* Seoul, Korea: Maumsarang.

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

Harkness, A. R., McNulty, J. L., & Ben-Porath, Y. S. (1995). The Personality Psychopathology Five (PSY–5): Constructs and MMPI–2 scales. *Psychological Assessment, 7*, 104–114.

Kim, I. J., Han, K., Lim, J., Lee, J., Min, B., & Moon, K. (2005). *Korean MMPI-2 user manual.* Seoul, Korea: Maumsarang.

Kim, B. S. K. & Omizo, M. M. (2003). Asian cultural values, attitudes toward seeking professional psychological help, and willingness to see a counselor. *The Counseling Psychologist*, *31*, 343-361. doi: 10.1177/0011000003031003008

Kim, D., Wampold, B. E., & Bolt, D. M. (2006). Therapist effects in psychotherapy: A random-effects modeling of the national institute of mental health treatment of depression collaborative research program data. *Psychotherapy Research, 16*(2), 161-172. doi:http://dx.doi.org/10.1080/10503300500264911

Lutz, W., Leon, S. C., Martinovich, Z., Lyons, J. S., & Stiles, W. B. (2007). Therapist effects in outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling Psychology, 54*(1), 32-39. doi:http://dx.doi.org/10.1037/0022-0167.54.1.32

Masuda, A., Suzumura, K., Beauchamp, K. L., Howells, G. N., & Clay, C. (2005). United States and Japanese college students' attitudes toward seeking professional psychological help. *International Journal of Psychology, 40* (5), 303-313. doi: 10.1080/00207590444000339

Nichols, D.S. (2006). The trials of separating bath water from baby: A review and critique of the MMPI–2 Restructured Clinical scales. *Journal of Personality Assessment, 87*, 121–138.

Patrick, C. J., Curtin, J. J., & Tellegen, A. (2002). Development and validation of a brief form of the Multidimensional Personality Questionnaire. *Psychological Assessment, 14*, 150 –163. doi:10.1037/1040-3590.14.2.150

Rouse, S. V., Greene, R. L., Butcher, J. N., Nichols, D. S., & Williams, C. L. (2008).What do the MMPI–2 Restructured Clinical scales reliably measure? Answers from multiple research settings. *Journal of Personality Assessment, 90*, 435–442.

Sellbom, M. & Ben-Porath, Y. S. (2005). Mapping the MMPI–2 Restructured Clinical Scales onto normal personality traits: Evidence of construct validity. *Journal of Personality Assessment, 85* (2), 179-187.

Sellbom, M., Ben-Porath, Y. S., & Graham, J. R. (2006). Correlates of the MMPI–2 Restructured Clinical (RC) scales in a college counseling setting. *Journal of Personality Assessment, 86*, 89–99.

Sellbom, M., Graham, J. R., & Schenk, P. W. (2006). Incremental validity of the MMPI-2 Restructured Clinical (RC) Scales in a private practice sample. *Journal of Personality Assessment, (86)*2, 196-205.

Simms, L. J., Casillas, A., Clark, L. A., Watson, D., & Doebbeling, B. N. (2005). Psychometric evaluation of the Restructured Clinical Scales of the MMPI-2. *Psychological Assessment, (17)*3, 345-358.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*, 245-251.

Stevens, J. P. (1984). Outliers and influential data points in regression analysis. *Psychological Bulletin, 95*(2), 334-344. doi: 10.1037/0033-2909.95.2.334

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics* (5th ed.). Boston, MA: Pearson Education.

Tellegen, A., & Ben-Porath, Y. S. (2008). *MMPI-2-RF (Minnesota Multiphasic Personality Inventory—2 Restructured Form): Technical manual.* Minneapolis, MN: University of Minnesota Press.

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

University of Minnesota Press. (n.d.). *Available Translations: MMPI-2, MMPI-2-RF, MMPI-A.* Retrieved April 22, 2013, from http://www.upress.umn.edu/tests/MMPI_Translations.html

Wampold, B. E., & Brown, G. S. (. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73*(5), 914-923. doi:http://dx.doi.org/10.1037/0022-006X.73.5.914

Wetzler, S., Khadivi, A., & Moser, R. K. (1998). The use of the MMPI-2 for the assessment of depressive and psychotic disorders. *Assessment, 5*(3), 249-261. doi:10.1177/107319119800500305

Zhang, N. & Dixon, D. N. (2003). Acculturation and attitudes of Asian international students toward seeking psychological help. *Journal of Multicultural Counseling and Development, 31,* 205-222.