Effects of Random Responding on the Interpretability of the MMPI-2-RF Substantive Scale Scores

A dissertation submitted to Kent State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

by

Wendy R. Dragon

August, 2012
Dissertation written by
Wendy R. Dragon
B.A., Kent State University, 2003
M.A., Kent State University, 2008
Ph.D., Kent State University, 2012

Approved by

Yossef Ben-Porath, PhD, Chair, Doctoral Dissertation Committee
Jack Graham, PhD, Members, Doctoral Dissertation Committee
Benjamin Newberry, PhD,
Christopher Was, PhD,
Timothy Gallagher, PhD,

Accepted by

Maria Zaragoza, PhD, Chair, Department of Psychology
John R.D. Stalvey, PhD, Dean, College of Arts and Sciences
Table of Contents

I. List of Tables.......................................................................................................................v.

II. Acknowledgements........................................................................................................vii.

III. Introduction......................................................................................................................1
   Early validity scales designed to detect random responding.................................3
   Subsequent efforts to detect random responding......................................................5
      MMPI-2..........................................................................................................................5
      MMPI-A........................................................................................................................11
      MMPI-2-RF....................................................................................................................11
   Detecting test-taker-generated random responding..............................................14
   Effects of random responding on substantive scale scores.................................17
   Effects of random responding on criterion validity............................................19
      MMPI studies..............................................................................................................24
   Summary.........................................................................................................................28
   Research questions....................................................................................................30

IV. Methods.......................................................................................................................34
   Samples.........................................................................................................................34
      College Sample...........................................................................................................34
      Outpatient Sample.................................................................................................37
      Inpatient Sample.................................................................................................38
List of Tables

Table 1. Demographic characteristics of total sample and sample with valid MMPI-2-RF ................................................................. 35

Table 2. College sample criterion measures with their internal consistencies .......................... 42

Table 3. Outpatient sample criterion measures with their internal consistencies. .................. 45

Table 4. Civil forensic sample criterion measures ......................................................... 49

Table 5. The Effects of 30\%, 60\% and 90\% Random Response Insertion on MMPI-2 RF Scale Validity Coefficients in a College Sample ........... 53

Table 6. The Effects of 30\%, 60\% and 90\% Random Response Insertion on MMPI-2 RF Scale Validity Coefficients in an Outpatient Sample ...... 58

Table 7. The Effects of 30, 60, and 90\% Random Response Insertion on MMPI-2 RF Scale Validity Coefficients in an Inpatient Sample .......... 62

Table 8. The Effects of 30\%, 60\% and 90\% Random Response Insertion on MMPI-2 RF Scale Validity Coefficients in a Criminal Forensic Sample 65

Table 9. The Effects of 30\%, 60\% and 90\% Random Response Insertion on MMPI-2 RF Scale Validity Coefficients in a Civil Forensic Sample ... 67

Table 10. The Effects of 30\%, 60\%, and 90\% Random Response Insertion on MMPI-2-RF Scale Scores in a College Sample................................. 71

Table 11. The Effects of 30\%, 60\%, and 90\% Random Response Insertion on MMPI-2-RF Scale Scores in an Outpatient Sample......................... 76

Table 12. The Effects of 30\%, 60\%, and 90\% Random Response Insertion on MMPI-2-RF Scale Scores in an Inpatient Sample ......................... 81
Table 13. The Effects of 30%, 60%, and 90% Random Response Insertion on MMPI-2-RF Scale Scores in a Criminal Forensic Sample ............... 86
Table 14. The Effects of 30%, 60%, and 90% Random Response Insertion on MMPI-2-RF Scale Scores in a Civil Forensic Sample .................... 90
Table 15. The Effects of 30%, 60%, and 90% Random Response Insertion on the Pattern of MMPI-2 RF Higher Order Scale Elevations.............. 95
Acknowledgements

I would like to take this opportunity to acknowledge my academic advisor, Yossef Ben-Porath, whose guidance and dedication made it possible for me to complete this project. I would like to send my love and gratitude to my husband, Scott, whose patience and support continue to help me in more ways than I can count; my children, Samantha and William, who are my motivation and my life; and my father, who instilled in me the desire to strive for excellence. In addition, I believe it was my mother’s love and encouragement that has given me the strength to achieve this dream. I wish that she were here to see it.
Effects of Random Responding on the Interpretability of MMPI-2-RF Substantive Scale Scores

Psychologists have known since the early days of formal psychological assessment that a test-taker may provide inaccurate responses on a psychological test for a variety of reasons including carelessness, inability to respond to test stimuli, or a desire to present him or herself in a particular light. Consequently, the validity of a psychological test protocol must be evaluated and established before it is used to make decisions about the test-taker (Ben-Porath & Waller, 1992). Researchers have identified two broad types of protocol validity concerns: content-based invalid responding and non-content-based invalid responding (Ben-Porath, 2003). Content-based invalid responding occurs when a test-taker responds to item content in a way that generates an implausible picture of his or her psychological functioning. In contrast, non-content-based invalid responding occurs when a test-taker responds to items without regard for their content.

Historically, test authors, researchers, and users have focused on detecting content-based invalid responding and have devoted considerably less effort to the development and research of non-content-based invalid responding scales (Otto, 2008). However, for a test user to conclude that content-based invalid responding has occurred, he or she must first rule out non-content-based invalid responding, as non-content-based response styles can affect scores on content-based validity scales (Burchett, Dragon, Ben-
Porath, & Handel, 2010). In other words, detection of content-based invalid responding requires an initial assessment of non-content-based invalid responding.

Non-content-based invalid responding is divided into three major types: nonresponding, fixed responding, and random responding (Ben-Porath, 2003). Nonresponding occurs when a test-taker fails to answer an item or provides more than one answer, rendering the response unscorable. Fixed responding occurs when a test-taker answers test items in a scorable fashion, but prefers one answer option (e.g., “True”) over others regardless of item content. Random responding occurs when a test-taker answers items variably (e.g., arbitrarily choosing any response option) regardless of their content. Random responding is seen as a “problematic response set that can invalidate the results of a self-report inventory” (p. 139; Baer, Ballenger, Berry, & Wetter, 1997), and may be the hardest type of non-content-based invalid responding to detect, as well. Nonresponding is easily seen in a protocol. Test users often look for unscorable responses when test-takers turn in their answer sheets, and encourage them to revisit unanswered items or ones with multiple response options endorsed. Similarly, with extreme amounts of fixed responding, a casual inspection of the protocol will show the test-taker has engaged in a non-content-based response set. But even at extreme levels of random responding, a casual inspection of the protocol may not alert the test user to the presence of this response style. Yet many test-takers engage in at least partial random responding when taking a psychological inventory (Baer, et al., 1997; Berry, Wetter, Baer, Larsen, Clark, & Monroe, 1992).
Partial random responding describes a response style where test-takers shift between content-consistent and random response styles within a single instrument. Partial random responding may range from minimal (i.e., only a few items are responded to randomly) to extensive (i.e., almost all responses are random) and takes one of three forms. **Front random responding** occurs when test-takers respond randomly to initial test items but begin responding relevantly to items later in the test. **Back random responding** occurs when a test-taker changes from content-based responding in the beginning of a test to random responding later in the test. **Intermittent random responding** occurs when test-takers switch between random and content-based responding two or more times within a test (MMPI-2-RF; Clark, Girona, & Young, 2003).

A great deal of research has been published regarding random responding on the Minnesota Multiphasic Personality Inventory (MMPI) family of instruments. However, most of this research focuses on the creation and validation of scales to detect random responders, rather than on the interpretability (or validity) of substantive scale scores when random responding occurs.

**Early Validity Scales Designed to Detect Random Responding**

The developers of the MMPI (Hathaway & McKinley, 1943) created the first infrequent response scale in order to detect problematic protocols (the Infrequency, or F scale; Meehl & Hathaway, 1946). The F Scale was composed of items endorsed by less than ten percent of the normative sample of the MMPI. Because test-takers were unlikely to endorse a large number of low base rate items that covered a variety of content
domains, those who answered a large number of F items were seen as not responding to test content (Meehl & Hathaway, 1946). Early MMPI research revealed that the F scale was not only sensitive to random responding; it was also elevated when test-takers’ attempted to present themselves as having more psychological problems than they actually had (Meehl & Hathaway, 1946).

Concerned about the F scale’s lack of specificity to random responding, Buechley and Ball (1952) created the Tr Index scale to detect it. The Tr Index was composed of 14 pairs of duplicate MMPI items (repeated in the test booklet to facilitate machine scoring). An incongruent response to an item pair increased the scale score by one point. The purpose of the Tr Index was to separate high F scorers into random responders and those who responded validly but in a consistently bizarre fashion. A second scale, Carelessness (CLS; Greene, 1978), was created to assist users of the Tr index in discriminating between content-consistent and random responders. Instead of using duplicate item pairs, CLS was composed of psychologically-opposite item pairs. Early research showed that CLS and the Tr Index appeared to function similarly, classifying the same cases as random or content-consistent responders 92% of the time (Greene, 1978).

Although these early studies were illuminating, they did not provide test users with vital information about the ability of these scales to detect random responding. Later researchers (Nichols, Greene, & Schmolk, 1989; Rogers, Dolmetsch, & Cavanaugh, 1983) used computer-generated random response protocols to develop cut off scores and interpretive guidelines that maximized classification between random and content-consistent responders. They, like Greene (1978), reported that the Tr Index plus CLS
consistently produced effectiveness rates between 80 and 90 percent.

**Subsequent Efforts to Detect Random Responding.**

**MMPI-2.**

Although F was largely maintained in the restandardized version of the MMPI (the MMPI-2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989), new validity measures, the Back Infrequency (Fb), Variable Response Inconsistency (VRIN), and True Response Inconsistency (TRIN) scales were added. Fb was designed to assess the validity of responses to items appearing in the latter half of the MMPI-2. VRIN was developed in order to detect whether a test-taker answered the inventory items in a random manner. The scale is composed of 67 pairs of items that are either similar or opposite in content. TRIN was created to detect fixed responding, and is composed of 20 pairs of items that are opposite in content.

Several authors have investigated the ability to detect random responders using these and other scales. For example, Paolo and Ryan (1992) examined the classification accuracy of MMPI-2 scales and indices (F, Fb, VRIN[T-score], VRIN[raw], |F-Fb|, and F+Fb+|F-Fb|) at different cut scores in detecting random responding in a veteran sample. This sample was composed of mainly of inpatients (chemical abuse, 92%; psychiatry, 5%), although three percent of the sample came from outpatient clinics or civil forensic settings. Additionally, two hundred protocols were generated by computer to simulate random responding. VRIN, F, and Fb differentiated random from actual protocols, and a raw score of 13 on VRIN reasonably identified completely random protocols.
(effectiveness rate of 92%). |F-Fb|, and F+Fb+|F-Fb| were not as effective at detecting completely random response sets, and did not add incrementally to the standard validity scales (Paolo & Ryan, 1992).

Focusing on the ability to differentiate between back random responders and content-consistent responders, a later study also examined a number of standard validity scales and their derivative indices (F, Fb, VRIN, |F-Fb|, VRIN+|F-Fb| and F+Fb+|F-Fb|) in a college student sample (Cramer, 1995). The students completed the entire inventory, and then the researcher replaced the items at the end of the test with random responses, starting at a point 25, 50, or 75 percent of the way through the test. One fourth of the protocols had no items replaced with random responses and served as a comparison group. An additional all-random group was composed of computer-generated random responses. All of the indices (except F and |F-Fb|) reliably discriminated between content-consistent responders and all levels of random responders. Although Fb and F+Fb+|F-Fb| misclassified the fewest protocols, VRIN misclassified fewer protocols as levels of random responding increased (Cramer, 1995).

Clark, Gironda, and Young (2003) also examined the ability of the validity scales to detect back random responding on the MMPI-2. Using the MMPI-2 normative sample and a veteran sample, they replaced from 0 to 550 items in increments of 50 items from the back to front of the MMPI-2 booklet. They found |F – Fb| \(\geq 20\) was most effective at discriminating between back random responders and content consistent responders. Next, they examined the incremental validity of |F-Fb|, Fb, and VRIN in detecting random responders and found |F-Fb| entered first in a hierarchical regression for the normative
samples, whereas Fb alone entered first in the clinical sample. Additionally, they reported that the MMPI-2 Validity Scales demonstrated greater sensitivity to back random responding than the Validity Scales of the Personality Assessment Inventory (PAI; Morey, 1991).

Similar analyses were conducted using the MMPI-2 in a custody evaluation sample by replacing part (35, 50, or 65%) of the items at the end of the protocol with random responses (Pinsoneault, 2007). F and VRIN were the most effective in detecting the all-random protocols, although VRIN only identified half of the 50% random protocols. Pinsoneault constructed a number of derivative VRIN indices, with items located in the front third (VRIN_a), the back third (VRIN_d), or the front or back two-thirds (VRIN_c and VRIN_b), of the MMPI-2 as well as indices covering the first half (VRIN_1) and back half (VRIN_2) of the inventory. He created 3 F (F_a, F_b, and F_c) subscales; each subscale was limited to one-third of the inventory. He also examined whether $|F_b - F| \geq 30$ was effective at discriminating between random and content-consistent responders. All subscales and indices were significantly different with large effect sizes at the expected levels of random responding.

The largest benefit of using the VRIN scale to detect random responding is its relative insensitivity to over-reporting. This was demonstrated by Wetter, Baer, Berry, Smith, and Larsen (1992) in a college sample. Students were asked to either complete the MMPI-2 under standard instruction, to answer randomly, to malinger a moderate psychological disturbance, or to malinger a severe psychological disturbance. Although F
and Fb were elevated under both random responding and under-reporting conditions, VRIN was only elevated under random responding conditions.

Some researchers have used scales other than the standard validity scales or their derivatives to detect random responding. For example, Sewell and Rogers (1994) tested the efficacy of a new 16-item measure, Inconsistent Response (IR), to screen for random responding in a sample of outpatients. IR was able to differentiate outpatient protocols from computer-generated random response protocols effectively (94% effectiveness rate). The authors suggested that IR might be most useful to screen the presence of random responding, as IR is easier to score than other random response detectors. However, the widespread use of computerized scoring means that most clinicians have no need for a screening scale, and it is unclear why clinicians would not choose to consult measures such as F and VRIN, which have a greater research base.

In contrast to the wide array of studies published concerning random responding on the MMPI family of instruments, only one study (Handel, Arnau, Archer, & Dandy, 2006) has focused on the detection of fixed responding. Handel and colleagues (2006) evaluated the ability of the MMPI-2 and MMPI-A TRIN scales (described below) to detect varying degrees of acquiescent and counter-acquiescent responding at different base rates (50% and 9.3%). The MMPI-2 TRIN scale was sensitive to both acquiescent and counter-acquiescent response insertion in both conditions, more so in the former. However, raw mean scores on TRIN were significantly different under the fixed response conditions than they were from those of the content-consistent sample, even when only ten percent of acquiescent responses were inserted. Similar findings were reported when
counter-acquiescent responses were inserted, indicating that TRIN is sensitive to both acquiescent and counter-acquiescent forms of fixed responding. Although it is clear that more research with TRIN could extend the findings of Handel and colleagues (2006), TRIN appears to be as useful at detecting fixed responding as VRIN is in detecting random responding.

Research has also established the ability of the validity scales of the adolescent version of the MMPI (Minnesota Multiphasic Personality Inventory – Adolescent [MMPI-A]; Butcher, Williams, Graham, Archer, Tellegen, Ben-Porath et al., 1992) to detect random and fixed responding. The MMPI-A, developed for use with adolescents ages 14 – 18, is shorter than both the MMPI and the MMPI-2. During the development of the MMPI-A, some MMPI items were reworded to better apply to an adolescent population and other items were added to deal with adolescent behaviors and problems. The MMPI-A also contains F, VRIN, and TRIN scales, constructed in the same way as those found on the MMPI-2. In addition, the F scale is divided into two parts, F1 and F2, each covering one half of the test booklet. Archer and Elkins (1999) investigated the ability of these scales to classify both inpatient and outpatient adolescents from computer-generated random protocols. Additionally, they used a 10:1 base rate in order to examine the positive and negative predictive powers of these scales in more naturally occurring rates of random responders, and found that higher cutoff scores were needed to correctly classify protocols in lower base rates of random responding. Although all of the scales examined were helpful in classifying random and content-consistent responders,
only F and VRIN added to the differentiation of the random responding from the clinical group in a stepwise discriminant function analysis.

Archer, Handel, Lynch, and Elkins (2002) extended these findings by examining the ability of several scales (VRIN, F, F1, F2, and the $|F1 – F2|$ Index) to detect partial (33%, 50%, and 67%) random responding in adolescent inpatient protocols at two base rates (10:1 and 1:1). They found that VRIN, F, and F2 were good at detecting higher rates of random responding (50 and 67%) and complete random responding, but gave mixed results under lower rates of random responding (33%).

Also addressing the issue of partial random responding on the MMPI-A, Pinsoneault (2005) replaced the content-consistent responses of delinquent adolescents with 5 different levels (25, 33, 50, 67, or 75%) of random responses. He examined the ability of F1 and F2 to detect back random responding, and created a number of derivative VRIN indices where items were located in the front third (VRIN$_a$), the back third (VRIN$_d$), or the front or back two-thirds (VRIN$_c$ and VRIN$_b$), of the inventory. Additionally, he created three F ($F_a$, $F_b$, and $F_c$) subscales, which were limited to the first, second, or final third of the inventory, respectively. $F2 – F1$ (.95) and $F_c – (F_a + F_b)$ (.98) were most effective in detecting protocols with 50 to 75% and 25 to 33% random responses inserted, respectively. VRIN demonstrated the best sensitivity (.65) across all random responding conditions, and had good negative predictive power (1.00), positive predictive power (.92) and effectiveness (.92) for 50 to 75% random responders, but overall, failed to identify one-third of random responders.
**MMPI-A.**

The MMPI-A TRIN scale serves the same purpose as the MMPI-2 TRIN scale, and is constructed similarly. However, it has twenty-one item pairs, one more than its MMPI-2 counterpart. Only Handel and colleagues (2006) have examined its ability to detect fixed responding on the MMPI-A. Using the same manipulation that they used with the MMPI-2 TRIN scale, they created a data set with a base rate of 10:1 content-consistent versus non-content-based fixed responding protocols. Even in a low base rate sample, MMPI-A TRIN demonstrated high negative predictive power and relatively high positive predictive power, even at low levels of fixed response insertion, similar to the MMPI-2 TRIN scale.

**MMPI-2-RF.**

Recently, Ben-Porath and Tellegen (2008) created the Minnesota Multiphasic Personality Inventory – 2 Restructured Form (MMPI-2-RF) as an efficient yet comprehensive set of scales that assess personality constructs and symptoms of psychopathology. Due to the more limited item pool of the MMPI-2-RF, new protocol validity indicators had to be created to detect random responding. The random responding indicator for the MMPI-2-RF, Variable Response Inconsistency-Revised (VRIN-r), is composed of 53 pairs of items selected so that the members of each pair are similar in content. VRIN-r was created using a configural approach to the scoring of item pairs, such that one or both of the possible two configurations of items (i.e., “True-False”, “False-True”) can add a point to the scale in any given protocol (this process is described
in detail in Tellegen & Ben-Porath, 2008). They called these pair units c-composites; and a scorable configuration of items was called a positive score. The items comprising a c-composite had to be positively correlated; and when positively scored, c-composites had to be content-inconsistent. Additionally, in order to be placed on the scale, the positive score of a c-composite had to be statistically improbable. In developing the True Response Inconsistency –Revised scale (TRIN-r) for the MMPI-2-RF, Tellegen and Ben-Porath (2008) used a similar method to the one used to develop VRIN-r, except that c-composites for TRIN-r had to be negatively correlated and consist of “True-True” or “False-False” combinations.

Using the MMPI-2-RF normative sample and a clinical sample, Handel, Ben-Porath, Tellegen, and Archer (2010) evaluated the effects of increasing percentages of random responding on both VRIN-r and the MMPI-2 VRIN scales, by replacing increasing percentages of items with computer-generated random responses on the MMPI-2-RF at ten percent increments from zero to seventy percent. VRIN-r and VRIN mean scores were similar (within 1.5 T score points), and increased monotonically as greater percentages of random responses were inserted into protocols. Additionally, the percentages of protocols scoring above 80 on VRIN and VRIN-r had a similar, increasing pattern. Overall, the use of VRIN-r in the detection of random responding on the MMPI-2-RF was supported.

Although the MMPI-2-RF is shorter than the MMPI-2, test-takers may still engage in partial random responding. Subscales, each covering half of the items of the MMPI-2-RF, were constructed for the Infrequency – Revised (F-r) scale, the Infrequency
Pathology – Revised scale (Fp-r), and VRIN-r (e.g., VRIN-r1, VRIN-r2) scales to detect back random responding (Veltri, Dragon, & Ben-Porath; 2008). These scales were tested by substituting twenty, fifty, or eighty percent of items on the back half of the inventory with random responses in community and inpatient samples. In order to detect back random responding, difference scores (e.g., VRIN-r2 – VRIN-r1) were computed. The VRIN-r|2 – 1| Index demonstrated a medium effect size in both samples at 20% of random responses inserted. Step-wise regression analyses demonstrated that all indices added incrementally, although the F-r|2-1| index entered into the prediction equation first in the normative sample, and the Fp-r|2-1| index entered first in the clinical sample. These two studies support the ability of the validity indicators to detect both complete and some forms of partial random responding on the MMPI-2-RF.

Fixed responding is also a concern on the MMPI-2-RF; however, only one study to date has looked at the effects of fixed responding on validity and substantive scales. Handel and colleagues (MMPI-2-RF; 2010) artificially manipulated the amounts of fixed responding by inserting increasing percentages of acquiescent or counter-acquiescent responses at the individual protocol level. They reported that TRIN-r scores increased monotonically as greater levels of fixed responses (both acquiescent and counter-acquiescent) were inserted into the normative sample. The mean T scores for the scale became greater than 79T when only thirty percent of items were replaced with acquiescent responses or when forty percent of items were replaced with counter-acquiescent ones.
The studies just reviewed demonstrate both the utility of the MMPI instruments’ validity scales in detecting random and fixed responding and how computer-generated random protocols can be used to examine the utility of inconsistent response indicators. Nonetheless, some authors (e.g., Costa & McCrae, 1997) have suggested that computers create random data that are different from how test-takers would randomly respond. Studies using test-taker random responders are reviewed next.

**Detecting Test-taker-generated Random Responding**

To date, only six studies have looked at the utility of validity scales to detect test-taker-generated random responding in the MMPI family of instruments. In the first study to examine random responding using actual test-takers, Berry and colleagues (Berry et al., 1991) administered the test under standard instructions, but removed the test booklet and asked test-takers to randomly answer different amounts of items at the end of the test. A discriminant function analysis indicated an overall effectiveness rate of 92%, with 91% of the random and 96% of the fully valid protocols correctly classified. Berry and colleagues found that F, Fb, and VRIN were sensitive to presence of random responding.

Another study examined the utility of F, Fb, and VRIN in detecting self-reported partially random responding in college students, community volunteers, and job applicants (MMPI-2; Berry, Wetter, Baer, Larsen, Clark, & Monroe, 1992). These validity scales were reliably and positively correlated with self-estimated random responses in every sample but the job applicants. The authors suggested that the lower levels of random responding in the job applicant sample attenuated the correlations
between validity scales and the number of self-reported random responses. Test-takers were more likely to respond randomly toward the end of the test; however, test-takers also scattered random responses throughout the test.

Gallen and Berry (MMPI-2; 1996) examined the utility of the raw score of VRIN VRIN(raw) and VRIN(raw) +|F-Fb(raw)| to detect test-taker-generated back random responding at 6 levels (0, 100, 200, 300, 370, 400, or 500 valid items) in a college sample. Both measures discriminated between back random responders and content-consistent responders. Although VRIN(raw) +|F-Fb| had the highest effectiveness rate, VRIN alone also provided highly accurate classification.

Using the same college sample and manipulation as the study just described (Gallen & Berry, 1996) Gallen and Berry (1997) examined the utility of F, Fb, F and Fb derivative indicators(|F - Fb|, F+Fb+|F-Fb|), VRIN and VRIN subscales (VRINf, VRINb, and VRINm), VRIN+|F-Fb|, and IR (Sewell & Rogers, 1994) to detect random responding. They also examined the Blocked Infrequency Index (BI), which determines the percentage of F and Fb items endorsed in the scored direction within 100-item blocks on the MMPI-2 (with the sixth block composed of items 501 to 567). They determined that percentages above 27% accurately classified random from valid responders within each block. Although all indices differed significantly between protocols with and without random responding, VRINf, which was composed of item pairs on the VRIN scale if both items occurred before item 370, demonstrated highly accurate classification rates and is likely to have the parent scale's insensitivity to over-reporting. Additionally,
F, F+Fb+|F-Fb|, and BI had the highest hit rates across all examined base rates of random responding.

Only two studies have examined the ability of validity scales to detect actual adolescent random responders. Baer, Ballenger, Berry, and Wetter (1997) used a methodology similar to the one used by Berry and colleagues (1991) to examine the ability of F, F1, F2, and VRIN in detecting partial back responders on the MMPI-A. All indicators were sensitive to partial and totally random protocols; except for F1, which was only effective in discriminating content-consistent protocols from those that engaged in random responding during the first half of the test. Although all three F scales had higher effectiveness rates, elevations, and effect sizes, VRIN demonstrated high classification rates, large effect sizes, and was purported to be insensitive to over-reporting.

In an attempt to determine whether the MMPI-A VRIN scale was insensitive to over-reporting, Baer, Kroll, Rinaldo, and Ballenger (1999) instructed nonclinical adolescents to over-report or randomly respond, and instructed a clinical group to respond content-consistently or randomly to the test. VRIN was only sensitive to random responding; whereas the F scale and F-K were sensitive to both random responding and over-reporting.

Overall, these studies demonstrate that the MMPI instruments’ validity scales are able to detect random responses, whether they are generated by test takers or computers. An inherent assumption in this research is that scales and tests are less useful or accurate if responded to randomly than they are if responded to in a content-consistent manner. In
other words, it is assumed that random responding has a negative effect on substantive scales’ criterion validity (otherwise, why detect it?). Only a small body of research has put this assumption to an empirical test.

**Effects of Random Responding on Substantive Scale Scores**

Many studies have demonstrated that mean scores on substantive scales are significantly different in high versus low random responding groups. On the original MMPI, content-consistent and computer-generated random profiles produced significantly different means on all of the Clinical Scales, F, and the L Scale (Rogers, Dolmetsch, & Cavanaugh, 1983).

Differences in MMPI-2 Clinical Scale means for random responding and content-consistent responders have also been reported. Clark and colleagues (2003) used differences in mean scores greater than 5T score points as a benchmark for gauging clinically significant differences. The Clinical Scales were resistant to the effects of back random responding until 250 items or more at the end of the test were replaced with random responses (due to the fact that no items on the Clinical Scales occur past item 370, virtually no items could be affected until more than 200 items were replaced with random responses). The mean scores of six out of 15 Content Scales were affected when only 100 back random responses were inserted, and all of the substantive scales were more than 5T score points different from content-consistent scales when more than 400 back random responses were inserted. Wetter and colleagues (1992) also reported that random responders had higher scores than content-consistent responders on all of the
Clinical Scales, although individuals faking severe mental illness produced higher scores than either content-consistent or random responders.

On the MMPI-A, only scores on Clinical Scales 7 and 9 were reportedly unaffected by extreme random responding; whereas scores on Clinical Scales 6 and 8 were distorted even when random responding was as low as twenty-five percent (Baer et al., 1997). Archer and Elkins (1999) found that among the Clinical Scales on the MMPI-A, only Scale 7 was unaffected by random responding when comparing computer generated profiles with a clinical adolescent population. Another study demonstrated mean differences on the MMPI-A Clinical Scales between over-reporters, content-consistent responders, clinical test-takers, and random responders in an adolescent population (Baer et al., 1999). In this study, random responders scored significantly higher than a clinical sample on three Clinical Scales (Scales 2, 3, and 5) and higher than content-consistent responders on seven of the ten Clinical Scales.

A recent study examined the effect of random responding on RC Scale scores using untruncated and unrounded T scores¹ (Handel, Ben-Porath, Tellegen, & Archer, 2010). They used the standard error of measurement (SEM) as a benchmark for

¹ Truncating refers to the process of limiting possible T score values to a particular range. In the case of the MMPI family of instruments, T scores are limited to values between 30 and 120T for the MMPI-2 (Butcher et al., 1989) and 20 and 120T for the MMPI-2-RF (Tellegen & Ben-Porath, 2008). Clinically, the interpretation of these extreme values would be the same with or without truncating. Rounding is the process by which numbers are made easier to work with by dropping unnecessary decimal places. Both truncating and rounding scores makes the scores easier to work with in a clinical setting, but introduces error into scales (Tellegen & Ben-Porath, 2008). Although the amount of error incorporated into any one protocol may be minimal; the accumulation of these errors in a sample introduces unnecessary error. Therefore this process should be avoided for research purposes.
determining whether substantive deviation from mean scores occurred due to random response insertion. SEMs ranged from 3 T score points (RCd) to 6 T score points (RC2, RC6, and RC8). RC6 was most sensitive to random responding, deviating from the original score by two SEMs after only 20% random response insertion. RC8 was affected similarly after only 30% random response insertion and RCd and RC1 were affected after 40% random response insertion. Using the 5T score point criterion for clinical significance, one scale’s scores were affected with as little as 10% (RC6) random response insertion, and other scales were affected at 20% (RC1 and RC8), 30% (RCd, RC2, and RC4), and 40% (RC7) random response insertion. At 70% random response insertion all scales except RC3 and RC9 exhibited highly distorted mean scores that would impact their interpretation.

**Effects of Random Responding on Criterion Validity**

Studies that examine substantive scales’ criterion validity have focused primarily on the effect random responding has on validity coefficients, although researchers have also used moderated regression to examine how random responding affects criterion validity. Because only three studies have looked at the effect of random responding on the criterion validity of MMPI instruments (two with the MMPI-2, and one with the MMPI-2-RF), studies that have examined this issue with other instruments are reviewed as well.

Hough, Eaton, Dunnette, Kamp, and McCloy (1990) examined the criterion validity of the Assessment of Background and Life Experiences (ABLE) using enlisted
military personnel who completed the test in one of three conditions: fake bad, fake good, or honest. They created a Nonrandom Response scale composed of 11 items with obvious correct responses. Lower scores on this scale were intended to indicate more random responding, although scores on this scale were also depressed in the fake-bad condition (3.04 standard deviations lower than the honest responders). Hough and colleagues split their sample into more and less random responders using this scale and reported that random responding moderated the criterion-related validities for the ABLE content scales. When correlations for content-consistent and random responders were compared, more than half (19 of 33) of the validity coefficients were significantly lower in the random group.

Costa and McCrae (1997) divided a community sample evenly into groups of high, medium, and low consistency based on an inconsistency index using pairs of highly correlated items (originally used in Costa & McCrae, 1992). Convergent validity correlations were computed between the five Revised NEO Personality Inventory (NEO–PI–R) domains and Goldberg’s Big Five (1992) adjective markers. Across the five trait factors, they reported that only one exhibited an appreciable reduction in convergent validity with greater levels of response inconsistency, and that even the least consistent 15% of the sample produced mean validity coefficients comparable to those of the most consistent test-takers. However, this study has significant limitations. For example, it is impossible to determine from the reported findings how inconsistent the low consistency group was, as no validation of the consistency measure was reported. Additionally, the authors did not provide information needed to evaluate or replicate their results.
Furthermore, correlations with extra-test criteria may not have declined significantly due to a restricted range of random responding. Furthermore, contrary to their interpretation of the results, examination of Costa and McCrae’s (1997) findings indicates that three of the five domain scales did demonstrate a modest decrease in validity between groups.

McCrae, Stone, Fagan, and Costa (1998) examined possible moderators of personality agreement between married couples. Along with personality, demographic characteristics, social perception variables, and relationship variables, random responding was tested in each spouse as one possible moderator of agreement. Random responding was measured with a semantic consistency scale, Inconsistency (INC; Schinka, Kinder, & Kremer, 1997). INC is a relatively brief scale consisting of ten item pairs; with five item pairs that are usually answered in the same direction and five that are usually answered in opposite directions. Decreases were observed in half of the correlations between random responding and a profile similarity index in the expected direction. The results of moderated regression analyses for scales predicting spouse ratings were largely non-significant; however, the researchers indicated they only had a power level of .20 to detect small effects. Furthermore, the authors recruited participants from a larger longitudinal study. Individuals who participated in this study are likely to be more invested in providing accurate responses than the average test-taker, given that they willingly volunteer their time for the larger study. This sample may have a lower than average base rate of random responding, which would serve to obscure the effects of any random responding on substantive scale scores.
Kurtz and Parrish (2001) hypothesized that if a semantic inconsistency scale was effective, then NEO-PI-R protocols identified as potentially invalid by INC would have reduced reliability and convergent validity across self-report and informant criteria. They compared low, medium, and highly inconsistent (random) responders’ protocols with information provided by friends or family. Regardless of level of inconsistency, contingency coefficients were the same for each subgroup. The authors concluded that correlations between self-report and informant data across conditions provided evidence that semantic consistency approaches to assessing protocol validity may overestimate the prevalence of random or careless response behavior in standard administration conditions. However, less than 20 percent (n = 14) of their sample was seen as highly inconsistent on a brief measure with limited research (one study) to back its utility at detecting totally random data, and no research had been conducted to determine whether it was effective under partially random conditions. They also did not address the fact that the cut score they used on INC for discriminating low from moderate random responders was lower than the mean score for the scale in the validation study (Schinka et al., 1997). The authors acknowledged that the results of these regression analyses remain somewhat ambiguous due to range restriction, and suggested that the impact of inconsistency may be greater at higher levels than lower ones. In fact, although not mentioned by the authors, all of the retest intraclass correlations (except one) were lower in the self-report high inconsistency group than in the low inconsistency group. And there were modestly lower convergent correlations in the moderate inconsistency group. The manner in which these results were interpreted led a few researchers to conclude that consistency versus
inconsistency may be an individual-differences variable, but one that does not influence protocol validity; however, this interpretation is not supported by the results of the study.

In addition to computing validity coefficients, Piedmont, McCrae, Riemann, and Angleitner (2000) conducted moderated regression analyses to examine the effect of random responding on the criterion validity of the NEO-PI-R and the Multidimensional Personality Questionnaire (MPQ). They reported that only 6% of regressions were significant for validity scale interaction and that the significant regressions showed no discernable pattern. Furthermore, they reported that correlations were often higher in the less valid group than in the more valid group. A number of methodological problems were evident with this study. They used a relatively small sample size (178 people) composed of two distinct populations: graduate pastoral students and college undergraduates in their first study. The researchers did not offer a coherent rationale as to why they thought it would be possible to examine the effect of all content and non-content-based invalidity measures in a single aggregated analysis. They tested the entire set of validity scales by designating as aberrant any test-taker who scored in the top 5% on any one of the scales without discussing previously validated cutoff scores on any of those measures. Because they combined test-takers who are elevated on scales measuring radically different response styles (e.g., random responding, acquiescence, and social desirability), it remains unclear whether poorly constructed validity scales concealed the utility of more effective ones or whether other clinical cutoffs on these measures would have been effective. Indeed, the authors stated that the author of the MPQ only recommended using two of its Validity Scales, yet they used four additional MPQ
Validity Scales in their analyses. Across both studies, they noted that one scale designed to be sensitive to random responding, the Variable Response Inconsistency scale (VRIN; Multidimensional Personality Questionnaire, Tellegen, 1982), showed promise – as increases in VRIN led to decreases in validity coefficients in 27 of 36 comparisons. Nevertheless, these authors suggested that clinicians should limit the use of validity scales.

Johnson (2005), using a different methodology, sought to examine the effect of high versus low response consistency on factor loadings on a proxy measure developed to mimic the NEO (IPIP-NEO; International Personality Item Pool representation of the NEO-PI-R; Goldberg, 1999). The author constructed two consistency measures (one using a form of split-half reliability, the other using psychometric antonyms) and then conducted separate item-level principal component factor analyses using scorers from the upper and lower quartiles on both measures. High scorers did not produce a clearer factor structure than low scorers. One significant methodological concern should be noted; neither validity scale was tested to see if it was effective at discriminating between content-consistent and random responders. Because these scales may not differentiate between random and content-consistent responders, it is impossible to interpret these results as anything but support that these validity scales lack utility in detecting random responding.

**MMPI studies.**

As stated earlier, only three studies have looked at the effects of random responding on the MMPI family of instruments. Archer, Fontaine, and McCrae (1998)
correlated MMPI–2 Clinical Scale scores with Revised Symptom Checklist–90 scores (SCL-R; Derogatis, 1983) in a sample of psychiatric inpatients. High VRIN scorers (patients who had VRIN scores greater than or equal to 80T) had lower validity coefficients than low VRIN scorers in 206 of the 216 comparisons, and correlations were significantly lower in 67 comparisons. These authors noted that because of the frequency of significant correlations in the elevated VRIN scale subsample, it is unlikely that these protocols were the result of completely random data. Nevertheless, higher response inconsistency was clearly associated with lower predictive validity on substantive scales.

Subsequently, McGrath, Rashid, Hayman, and Pogge (2002) examined whether protocol validity influenced correlations between MMPI-2 code types and criterion measures. They found that the average correlations between MMPI-2 code types and clinician data declined by .02 after excluding cases because of elevated scores on any response bias indicator (F, TRIN, VRIN, or Cannot Say). Unfortunately, their study had a number of methodological limitations. For example, they created two groups: valid and invalid. Protocols with an elevated score on any one of the four validity indicators were placed in the invalid group. Using only two groups, it is impossible to know if one or more of these indicators were effective. In one of their comparisons the correlations that they were working with was small -- although individual correlations were not reported, mean correlations of code types with criterion measures was .12. Because the criterion validities in this group were so low to begin with, there was not much room for loss of validity. Another limitation was the small sample sizes that were used (62 in the 6-8/8-6 code type group, and 65 in the 7-8/8-7 code group). Additionally, they only examined
this hypothesis using two code types, 6-8/8-6 and 7-8/8-7; other code types or single scales may react differently. Finally, they did not report on whether substantive scale scores differed between the two groups. Even if validity was not moderated, scale score distortion may have prevented a test user from determining a code type at all, much less the same one that was classified based on a content-consistent protocol.

Recently Handel and colleagues (2010) evaluated the effects of increasing degrees of non-content-based invalid responding on convergent and discriminant validity coefficients for the Restructured Clinical (RC; Tellegen, Ben-Porath, McNulty, Arbisi, Graham, & Kaemmer, 2003) scales of the MMPI-2-RF. They correlated each RC scale with one representative convergent and one discriminant criterion taken from the 18-item version of a clinician rating instrument, the Brief Psychiatric Rating Scale (BPRS; Overall & Gorham, 1962). The BPRS lacks appropriate criterion measures for RC3 and RC9, so those scales were not examined. All RC Scales’ validity coefficients were unaffected at levels of less than 30% random response insertion; but modest degradation of validity coefficients was evident at 30% random response insertion for RC1 (4% less criterion variance accounted for), RC2 (5% less variance) and RC8 (5% less variance). In the 70% random response insertion condition, the percentage of variance accounted for was between four (RC7) and 11% (RC2) lower than in the content-consistent condition. These results indicated that the RC Scales validity coefficients are relatively robust in the face of mild and moderate degrees of random responding, but do begin to deflate as the percentage of random responding increases.
In their review of response styles as a source of bias, McGrath, Mitchell, Kim, and Hough (2010) examined studies across four domains of assessment: personality and psychopathology, employment, forensic, and eligibility for disability. They pointed out that the use of validity scales assumes that response bias suppresses or moderates the criterion-related validity of substantive scales and that those validity scales are capable of detecting response bias. They computed bias by looking at mean correlations across studies that focused on criterion-related validity of indicators said to detect inaccurate responding, and reported a slightly higher mean validity coefficient across invalid samples than across valid ones (.29 vs. .27). Additionally, they reported that interaction terms, which look at increment in fit, showed almost no mean improvement across 35 comparisons. However, research has demonstrated that even small moderator effects can be important (McClelland & Judd, 1993) and by using an aggregated analysis, small but meaningful effects for any of the constituent scales may have been obscured.

McGrath and colleagues did indicate that careless or random responding may represent a bias, but cautioned that those findings were predicated on a small set of studies. Even so, McGrath and colleagues did not support the use of validity measures, remarking “false positives have their cost” (p. 450). However, false negatives have their cost as well. Random responding may affect substantive scale interpretation through the distortion of scale scores, and ignoring the presence of random responding may lead to the interpretation of misleading scores that can result in unfortunate real-world implications.
For example, a randomly responding test-taker may obtain an elevation on a scale measuring antisocial behaviors. In a forensic setting, an erroneous report that the test taker has antisocial tendencies is more likely to lead to a negative outcome than a report that he or she engaged in random responding. Because of the likelihood of misinterpretation, the effect of random responding on substantive scale scores must be determined, and the appropriate use of random response measures may prevent test users from making inappropriate interpretations about test-takers. As indicated above, one of the biggest concerns that test users should have about interpreting substantive scale scores in the presence of random responding is the degree to which those scores are distorted.

**Summary**

The literature reviewed in this introduction suggests that random responding would affect the interpretability of psychological test scores both by modest degradation in the criterion validity of substantive scale scores, as well as by substantive scale score distortion. Researchers have examined the moderation of validity under random responding conditions amidst a number of other response sets which may obscure their results (e.g., Piedmont et al., 2001; McGrath et al., 2010). Other studies have used poorly validated or under-researched validity scales to examine whether validity is moderated by random responding, and then concluded that there is only minimal or no evidence of moderation using validity scales (e.g., Costa & McCrae, 1997; Johnson, 2005). Additionally, because research demonstrates that validity coefficients do not drop to zero when random responding is suspected, a few researchers (Costa & McCrae, 1997) have
concluded test users can interpret scale scores on apparently content-inconsistent protocols. Their position presupposes that a correlation between test scores and an extra-test criterion is sufficient evidence to prove it is appropriate to interpret seemingly random scores; however, substantive scales may exhibit robust correlations with extra-test criteria under partial random responding conditions and yet provide erroneous interpretive information due to score distortion.

As discussed earlier, random responding may significantly change test scores depending on the keying of items on a scale and their endorsement rate. On scales composed of items that a normative (typically a non-clinical) sample rarely endorsed in the keyed direction (as would be the case for most measures of psychopathology, some more so than others), a test-taker will reach an interpretive cutoff after endorsing only a few such items. A random responder has a 50% probability of answering any dichotomously-scored item (e.g., one that can be answered true or false) in the keyed direction. Thus, even a modest amount of random responding can significantly elevate standard scores beyond interpretable thresholds without substantially altering the scale’s criterion validity. Although random responses would not be correlated with the extra-test criterion, any content-consistent item responses (i.e., ones that the test-taker answered based on their content) will remain correlated with the extra-test criterion. The remaining content-consistently responded to items preserve a correlation between the scale and a criterion measure. At an extreme level of random responding, where very few content-consistent responses remain in a scale its correlation with extra-test measures will degrade to a point indicating complete loss of validity. However, because even a
relatively small number of items answered randomly in the keyed direction may
significantly inflate standard scale scores on measures of psychopathology (as just
reviewed), interpretation of such scores will be problematic before any evidence of
appreciable decrement in criterion validity will be evident.

Although, as reviewed here, distortion of scores on the substantive scales of the
MMPI family of instruments under random responding is well documented, no studies
have examined in detail the impact of score distortion under random responding
conditions on the interpretation of MMPI protocols. Moreover, only three studies have
examined the effects of random responding on the validity coefficients of these
instruments (Archer, Fontaine, & McCrae, 1998; Handel et al., 2010; McGrath, Hashid,
Hayman, & Pogge, 2002).

**Research Questions**

To more fully explore the effects of random responding on the interpretability of
the MMPI-2-RF substantive scales, the proposed study will address the following
research questions:

1. *Does prediction bias occur as a function of random responding?*

As random responding increases, do validity coefficients become increasingly
attenuated? To explore the effect of random responding on validity coefficients, this
study replicates and extends Handel and colleagues’ (2010) findings. Validity
coefficients are likely to remain relatively robust in the presence of mild (30%) and
moderate (60%) amounts of simulated random responding (random response insertion).
Such a finding would be consonant with previous research, and is likely to be related to the strength of the correlations between the measures. Because the correlation between two measures is not perfect to begin with, typically tending to hover in the .25 to .40 range, a great deal of variance (between 84 and 95% for correlations of this magnitude) in one measure is not accounted for by its association with the other. Increases in random responding will ultimately change the correlation between the two variables; however, the weaker the association between the predictor (MMPI-2-RF substantive scale) and criterion to begin with, the greater the amount of random responding that will be needed to make a noticeable difference in the correlation between scores on MMPI-2-RF scales and relevant criteria. Thus, scales with stronger correlates should show greater deflation in validity coefficients than weaker ones.

This study also differs from existing ones in that the effects of random responding on validity coefficients are studied in various settings, using different criterion measures across a number of samples, including college students, clinical outpatients, clinical inpatients, forensic evaluatees, and civil forensic evaluatees. Additionally, higher levels of random responding (up to 90%) than were examined in the Handel study are simulated, to determine whether extreme amounts of random responding exhibit a significantly greater effect on validity coefficients than mild and moderate levels of random responding. Finally, a greater array of MMPI-2-RF substantive scales are examined, including the Higher-Order (H-O), Restructured Clinical (RC), and Specific Problems (SP) Scales. By examining the effect of random responding on a wide variety of scales, it can be ascertained whether the breadth and length of scales moderate the effect of
random responding on validity coefficients. Longer scales may be more resistant to the
effects of random responding, as they tend to be more reliable. Since random responding
impacts reliability (and by extension criterion validity), a measure with higher initial
reliability may be more resistant to the effects of lower levels of random responding.

(2) **Does random responding distort scores on substantive scales in a manner that
would affect their interpretation?**

As just reviewed and discussed, random responding increases means and standard
deviations of substantive scales scores, in some cases to levels that would trigger
different interpretations. This study examines whether the effect of random responding is
more dramatic in some samples or on certain scales than it is in others. Examination of
means and standard deviations of substantive scale scores under varying levels of random
responding replicates and extends the findings of Handel and colleagues (2010), using a
greater variety of samples, more scales and including a higher level of random
responding.

The hypothesis that scales composed of items with particularly low endorsement
rates (such as RC6 or Suicidal Ideation [SUI]) should exhibit greater score distortion is
examined. It is expected that lower levels of random responding will be needed to
produce significant scale score distortion on these MMPI-2-RF scales.

In addition to looking for changes in mean scores and standard deviations as a
function or random responding, analyses designed to indicate whether patterns of
elevation change as a function of random responding are conducted. Analyses focus on
the H-O Scales, which demarcate clinically important individual variation in the basic
domains of affect (EID), thought (THD), and action (BXD; Ben-Porath & Tellegen, 2008). Interpretive recommendations in the test manual call for consideration of the relative elevations of these scales to determine the presence and relative salience of these broad domains of dysfunction for a given test taker. In examining these scales as a group, analyses focus on whether random responding creates elevations for content-consistent test-takers with no actual elevation on an H-O scale and whether individual scale score changes resulting from random responding alter the pattern of actual elevations on the H-O Scales. Due to the particularly low endorsement rates of THD items, it is expected that THD will exhibit the largest increase in scores as a function of random responding, thereby making it more likely that a test-taker appears to have thought dysfunction, rather than emotional or behavioral dysfunction, as random responding increases. At extreme levels of random responding, THD elevations are expected to overshadow scores on EID and BXD, making test-takers appear thought disordered on the MMPI-2-RF, when in fact they may primarily exhibit emotional or behavioral problems.
Methods

Samples

For this study, participants were drawn from five archival data sets, including a college sample, a clinical outpatient sample, a clinical inpatient sample, a forensic sample of criminal defendants, and a civil forensic sample. All of the samples have been used in previous research.

To effectively determine the effects of random response insertion, strict criteria were used to exclude protocols of individuals who may have engaged in actual random responding (VRIN < 70T). To minimize other threats to protocol validity, standard MMPI-2-RF inclusion criteria were used: Cannot Say (CNS) ≤ raw score of 15, True Response Inconsistency Scale (TRIN-r) < 80T, Infrequent Responses (F-r) < 120T, Infrequent Psychopathology Responses (Fp-r ) < 100T. Table 1 includes the number of participants excluded on these criteria, as well as basic demographic information for all full and final samples.

College sample.

A college sample was used in this study as a comparison sample. The college sample is composed of protocols completed by undergraduates who completed the MMPI-2 in exchange for credit in an introductory psychology course. They completed the MMPI-2 twice, along with a large number of criterion measures over two sessions with a one-week interval in between sessions. Students were randomly assigned to
Table 1. *Demographic characteristics of total sample and sample with valid MMPI-2-RF*

<table>
<thead>
<tr>
<th></th>
<th>College</th>
<th></th>
<th>Outpatient</th>
<th></th>
<th>Inpatient</th>
<th></th>
<th>Criminal forensic</th>
<th></th>
<th>Civil forensic</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Valid(^1)</td>
<td>Total</td>
<td>Valid(^1)</td>
<td>Total</td>
<td>Valid(^1)</td>
<td>Total</td>
<td>Valid(^1)</td>
<td>Total</td>
<td>Valid(^1)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>725</td>
<td>560</td>
<td>499</td>
<td>342</td>
<td>2208</td>
<td>1423</td>
<td>965</td>
<td>713</td>
<td>1102</td>
<td>746</td>
</tr>
<tr>
<td>Female</td>
<td>1036</td>
<td>904</td>
<td>716</td>
<td>558</td>
<td>717</td>
<td>490</td>
<td>348</td>
<td>282</td>
<td>1039</td>
<td>775</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>19.6</td>
<td>19.6</td>
<td>33.0</td>
<td>33.1</td>
<td>40.1</td>
<td>41.0</td>
<td>33.7</td>
<td>34.5</td>
<td>40.9</td>
<td>40.5</td>
</tr>
<tr>
<td>SD</td>
<td>3.3</td>
<td>3.2</td>
<td>10.2</td>
<td>10.0</td>
<td>14.0</td>
<td>14.4</td>
<td>11.1</td>
<td>11.3</td>
<td>10.9</td>
<td>11.0</td>
</tr>
<tr>
<td>Education (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>12.9</td>
<td>12.9</td>
<td>12.2</td>
<td>12.5</td>
<td>9.3</td>
<td>9.5</td>
<td>11.8</td>
<td>12.1</td>
<td>12.1</td>
<td>12.4</td>
</tr>
<tr>
<td>SD</td>
<td>1.3</td>
<td>1.3</td>
<td>2.2</td>
<td>2.2</td>
<td>3.4</td>
<td>3.2</td>
<td>2.3</td>
<td>2.2</td>
<td>2.7</td>
<td>2.5</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>86.2%</td>
<td>89.1%</td>
<td>76.3%</td>
<td>80.0%</td>
<td>78.6%</td>
<td>83.9%</td>
<td>77.2%</td>
<td>80.7%</td>
<td>--(^2)</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>8.7%</td>
<td>7.8%</td>
<td>21.1%</td>
<td>17.7%</td>
<td>15.4%</td>
<td>10.8%</td>
<td>21.7%</td>
<td>18.5%</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>African American</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>1.1%</td>
<td>1.2%</td>
<td>0.3%</td>
<td>0.3%</td>
<td>1.0%</td>
<td>0.6%</td>
<td>0.3%</td>
<td>0.2%</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Asian/ Pacific Islander</td>
<td>1.4%</td>
<td>1.1%</td>
<td>0.2%</td>
<td>0.3%</td>
<td>1.1%</td>
<td>1.1%</td>
<td>0.3%</td>
<td>0.3%</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Other</td>
<td>1.9%</td>
<td>1.8%</td>
<td>2.1%</td>
<td>1.6%</td>
<td>4.0%</td>
<td>3.6%</td>
<td>0.5%</td>
<td>0.3%</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Divorced</td>
<td>--</td>
<td>--</td>
<td>23.5%</td>
<td>24.3%</td>
<td>26.6%</td>
<td>26.5%</td>
<td>17.1%</td>
<td>17.8%</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Married</td>
<td>--</td>
<td>--</td>
<td>20.8%</td>
<td>20.2%</td>
<td>19.1%</td>
<td>19.8%</td>
<td>15.4%</td>
<td>16.5%</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Never married</td>
<td>--</td>
<td>--</td>
<td>41.1%</td>
<td>40.1%</td>
<td>43.8%</td>
<td>43.3%</td>
<td>35.9%</td>
<td>35.0%</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Separated</td>
<td>--</td>
<td>--</td>
<td>11.4%</td>
<td>12.5%</td>
<td>6.9%</td>
<td>7.0%</td>
<td>7.6%</td>
<td>5.3%</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Widowed</td>
<td>--</td>
<td>--</td>
<td>2.1%</td>
<td>2.1%</td>
<td>2.9%</td>
<td>2.9%</td>
<td>1.9%</td>
<td>1.9%</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

1. This column represents the number of individuals whose MMPI-2-RF protocols were valid in this sample using the given exclusion criteria. For some analyses, smaller samples may be used, as it may not be possible to obtain a desired criterion measure for everyone in a sample that produces a valid MMPI-2-RF.

2. Two dashes (--) indicate that insufficient or no information was available for these demographic characteristics.
complete the conventional version of the computerized MMPI-2 twice, or took the computerized conventional version once and an adaptive version at the other testing session\(^2\). The criterion measures were counterbalanced within each criterion packet, and the administration order of the criterion packets was similarly counterbalanced to avoid order effects. After excluding protocols based on the above criteria, the sample consists of 560 men and 904 women (total 1464). Participants had a mean age of 19.6 years (SD = 3.2) at the time of testing and are primarily European American (89.1%). Chi square analyses and \(t\)-tests were conducted to detect any demographic differences between included and excluded protocols in this sample, but none were found.

**Outpatient sample.**

This sample is composed of outpatients who completed the MMPI-2 after an intake interview at a Midwestern community mental health clinic. Participants were from varied economic backgrounds, and they were given a wide range of diagnoses, including depression, anxiety, substance abuse, and adjustment disorders. However, psychotic disorders were under-represented in this sample, as individuals with severe psychotic pathology were referred for services elsewhere. After excluding invalid protocols, the outpatient sample is composed of 342 men and 558 women for a total sample size of 900. Participants in the final sample were on average 33.1 years old (SD = 10.0) years old at the time of testing, and were primarily of European American ancestry (80.0 \%). No

\(^2\) For this study, only conventionally administered MMPI-2s were used. If an individual took the conventional version of the test twice, one of the resulting MMPI-2 s was chosen for inclusion in this study.
statistically significant differences were found between the samples’ ages, the proportion of one gender over another, or differences in the ethnic distribution between the groups. Participants who produced valid protocols reported that they had completed more education ($t(1214) = 3.101, p=.002$) than individuals who produced invalid protocols.

**Inpatient sample.**

The inpatient sample used in this study is composed of 1423 men and 490 women (1913 participants in total) who received psychiatric inpatient care at a large urban county medical center or at a Veterans’ Administration medical center located in the same metropolitan area. The sample is composed primarily of individuals who voluntarily admitted themselves for psychiatric treatment; however, a portion of the patients were being held, either for observation, by court order, or until transportation to another facility could be arranged. Individuals in this sample were given a variety of Axis I diagnoses including substance abuse or dependence; organic brain disease; psychotic disorders including schizophrenia; and mood disorders, such as anxiety, depression, or bipolar; Post-Traumatic Stress Disorder; and an assortment of other eating, somatization, or sexual disorders (Arbisi, Ben-Porath, & McNulty, 2003). Forty percent of this sample had been given multiple Axis I diagnoses. In addition, fifteen percent of this sample met criteria for an Axis II diagnosis. The single most frequent Axis II diagnosis was Personality Disorder, Not Otherwise Specified (8.5%); however, six percent of the sample met criteria for Antisocial Personality Disorder. Almost half of the sample (41%) had between one and three previous hospitalizations, and an additional thirty-nine percent
had no previous hospitalizations. This sample was primarily European American (83.9%), with a mean age of 41.0 years (SD = 14.4) at the time of testing. Participants who produced valid protocols reported that had slightly more education ($t(3824) = 2.142$, $p=.03$) than individuals who produced invalid protocols.

**Criminal forensic sample.**

This sample was collected at a Midwestern court clinic where individuals are routinely given psychological evaluations to aid the court in answering forensic questions such as competence to stand trial and sanity at the time of an illegal act, and has been used in previous research (Sellbom, Ben-Porath, & Stafford, 2007). In this setting, the MMPI-2 is routinely administered as part of court-ordered evaluations. There are 713 men and 282 women (for a total of 995) in the final forensic sample. The mean age of this sample was 33.7 years (SD 34.5). Participants who produced valid protocols were on average more educated ($t(1312) = 3.162$, $p<.01$) than individuals who produced invalid protocols.

**Civil forensic sample.**

This sample is composed of archival data collected from non-head-injury disability claimants and counseling clients attending private practice, and has been used in previous research (Gervais, Ben-Porath, Wygant, & Green, 2007). A number of diagnoses were present in this sample including chronic pain (39%) anxiety (23%), orthopedic diagnoses (17%) or depression (15%). The sample is comprised of 746 men and 775 women with a mean age of 40.5 years (SD = 11.0), who primarily identify
themselves as European American (80.7%). Participants who produced valid protocols were on average more educated ($t (2140) = 1.092, p<.05$) than individuals who produced invalid protocols.

**Instruments**

**MMPI-2-RF.**

The Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008) is a 338-item restructured version of the MMPI-2. It consists of a three-tiered hierarchical scale structure. At the highest level, scales and the Higher-Order (H-O) Scales provide a broad indication of the presence and relative prominence of problems in the domains of emotional, thought, and behavioral dysfunction. At mid-level, the Restructured Clinical (RC) Scales capture the major distinctive components of the original MMPI Clinical Scales. The third, most narrowly focused level of measurement includes the Specific Problems (SP) Scales designed to assess facets of the RC Scales that warrant separate assessment, or significant aspects of psychological functioning that are not captured by the other RC Scales. The MMPI-2 was administered to all five samples, and MMPI-2-RF scores were derived from MMPI-2 protocols. Tellegen and Ben-Porath (2008) and van der Heijden, Egger, and Derksen, (2010) have shown that MMPI-2-RF scores derived from the full MMPI-2 are equivalent to ones obtained from administration of the shorter booklet.

**Sample specific instruments.**
For the college sample, self-report criterion information was gathered for all participants over the course of two sessions no more than two weeks apart. Criterion measure content spanned a wide range of domains including: emotional disorders, personality functioning, family functioning, substance abuse, magical thinking, perceptual aberration, and behavioral control. A full listing of the measures that were administered to this sample, along with the internal consistency coefficient values for each instrument reported in Forbey and Ben-Porath (2007) for this sample can be found in Table 2.

Internal consistency coefficients ranged from .65 for the Machiavellianism IV (MACH-IV; Christie & Geis, 1970) to .89 for the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) with an average alpha of .78.

In the outpatient sample, an intake form and the Patient Description Form (PDF; Graham, Ben-Porath, & McNulty, 1999) were used to gather data about the patients. The intake form was completed by a trained worker after interviewing the patient. Information about demographics, mental health history, mental status, and diagnostic impressions are available on this measure. The PDF is a 188-item measure filled out by the patients’ psychotherapists and covers a range of personality characteristics and symptoms. Scales were rationally and statistically derived in order to assess 25 major content domains of the PDF (Graham et al., 1999). The content of these scales ranges across emotional, behavioral, thought disorder, and somatic pathology. Graham and colleagues 1999 reported that the internal consistency coefficients for the PDF scales ranged from .69 to .92 for men and from .72 to .93 for women. Additionally, the SCL-90-
Table 2. *College sample criterion measures with their internal consistencies*

<table>
<thead>
<tr>
<th>Criterion measure</th>
<th>Author</th>
<th>Construct</th>
<th>Alpha&lt;sup&gt;1&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barratt Impulsivity Scale—10 (BIS-10)</td>
<td>Barratt, 1985</td>
<td>Motor impulsivity</td>
<td>.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td>General impulsivity</td>
<td>.80</td>
</tr>
<tr>
<td>Beck Depression Inventory (BDI)</td>
<td>Beck, Ward, Mendelson, Mock, &amp; Erbaugh, 1961</td>
<td>Symptoms of depression</td>
<td>.89</td>
</tr>
<tr>
<td>Behavioral Inhibition/Activation System (BIS/BAS)</td>
<td>Carver &amp; White, 1994</td>
<td>Funseeking</td>
<td>.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inhibition</td>
<td>.74</td>
</tr>
<tr>
<td>Drug Abuse Screening Test (DAST)</td>
<td>Skinner, 1982</td>
<td>Drug use/abuse</td>
<td>.86</td>
</tr>
<tr>
<td>Family Functioning Scale (FFS)</td>
<td>Tavitan, Lubiner, Green, Grebstein, &amp; Velicer, 1987</td>
<td>Positive family functioning</td>
<td>.89</td>
</tr>
<tr>
<td>Fears Questionnaire (FQ)</td>
<td>Marks &amp; Mathews, 1979</td>
<td>Social phobia</td>
<td>.68</td>
</tr>
<tr>
<td>Internal State Scale (ISS)</td>
<td>Bauer, Crits-Cristoph, Ball, DeWees, McAllister, Alahi, et al., 1991</td>
<td>Hypomaniac activation</td>
<td>.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Depressive symptomatology</td>
<td>.74</td>
</tr>
<tr>
<td>Machiavellianism—IV (MACH-IV)</td>
<td>Christie &amp; Geis, 1970</td>
<td>Negative beliefs about others</td>
<td>.65</td>
</tr>
<tr>
<td>Magical Ideation Scale (MIS)</td>
<td>Eckblad &amp; Chapman, 1983</td>
<td>Magical thinking</td>
<td>.81</td>
</tr>
<tr>
<td>Michigan Alcohol Screening Test (MAST)</td>
<td>Selzer, 1971</td>
<td>Alcohol use/abuse</td>
<td>.72</td>
</tr>
<tr>
<td>Obsessive Compulsive Scale (OCS)</td>
<td>Gibb, Bailey, Best, &amp; Lambirth, 1983</td>
<td>Obsessiveness</td>
<td>.70</td>
</tr>
<tr>
<td>Test Name</td>
<td>Author(s)</td>
<td>Description</td>
<td>Reliability Coefficient</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>------------------------------------</td>
<td>---------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>Perceptual Aberration Scale (PAS)</td>
<td>Chapman, Chapman, &amp; Raulin, 1978</td>
<td>Perceptual abnormalities</td>
<td>0.88</td>
</tr>
<tr>
<td>Screener for Somatoform Disorders (SSD)</td>
<td>Janca, Burke, Isaac, Burke, Costa e Silva, Acuda, et al., 1995</td>
<td>Physical complaints</td>
<td>0.83</td>
</tr>
<tr>
<td>State Trait Personality Inventory (STPI)</td>
<td>Spielberger, 1979</td>
<td>Trait Anger</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trait Anxiety</td>
<td>0.88</td>
</tr>
</tbody>
</table>

*As reported for this sample in Forbey and Ben-Porath (2007)*
R (Derogatis, 1983) was administered to this sample. Internal consistency coefficients for these scales were reported by Derogatis (1983) range from .77 on Psychoticism to .90 on Depression with an average value of .84. A full listing of scales and internal consistency values for the PDF and SCL-90-R scales in this sample can be found in Table 3.

A Record Review Form (RRF; Arbisi, Ben-Porath, & McNulty, 2003) was developed to provide a standardized and reliable means of extracting clinical information from patients’ charts at the inpatient settings. The information extracted from the charts included treatment and medication information, problems at admission, current stressors, criminal justice system involvement, diagnoses, and disposition. In addition, information on a broad array of personality characteristics, as well as emotional, cognitive, and somatic symptoms was coded by a trained undergraduate research assistant after reviewing the patient’s intake interview. Five percent of the charts were rated by two research assistants, in order to maintain consistent interrater reliability and prevent rater drift. The Kappa statistics for the RRF variables ranged from a low of .62 to a high of .97, with a mean of .83.

Another Record Review Form (RRF) was developed to extract historical information from charts in the forensic sample and used by Petroskey, Ben-Porath, and Stafford (2003). Information gathered in this sample included: social, romantic relationships, and parenting histories; educational and employment history; medical, psychiatric, and substance abuse histories; as well as legal and abuse history. Mental
Table 3. *Outpatient sample criterion measures with their internal consistencies*

<table>
<thead>
<tr>
<th>Criterion measure</th>
<th>Author(s)</th>
<th>Construct</th>
<th>Alpha</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient Description Form Scales</td>
<td>Graham, Ben-Porath, &amp; McNulty, 1999</td>
<td>Achievement Oriented</td>
<td>.84</td>
<td>.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aggressive</td>
<td>.82</td>
<td>.77</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Agitated</td>
<td>.69</td>
<td>.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Angry Resentment</td>
<td>.92</td>
<td>.90</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Antisocial</td>
<td>.90</td>
<td>.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Anxious</td>
<td>.87</td>
<td>.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Critical/Argumentative</td>
<td>.91</td>
<td>.90</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Defensive</td>
<td>.89</td>
<td>.89</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Depressed</td>
<td>.86</td>
<td>.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Emotionally Controlled</td>
<td>.88</td>
<td>.85</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Family Problems</td>
<td>.89</td>
<td>.90</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Histrionic</strong></td>
<td>.89</td>
<td>.90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Insecure</strong></td>
<td>.87</td>
<td>.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Introverted</strong></td>
<td>.84</td>
<td>.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Narcissistic</strong></td>
<td>.92</td>
<td>.93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Negative Treatment</strong></td>
<td>.89</td>
<td>.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Attitudes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Obsessive-Compulsive</strong></td>
<td>.83</td>
<td>.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Passive-Submissive</strong></td>
<td>.85</td>
<td>.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pessimistic</strong></td>
<td>.73</td>
<td>.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Procrastinates</strong></td>
<td>.75</td>
<td>.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Psychotic Symptoms</strong></td>
<td>.80</td>
<td>.87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Somatic Symptoms</strong></td>
<td>.92</td>
<td>.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stereotypic Masculine</strong></td>
<td>.92</td>
<td>.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Interests</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Suspicious</strong></td>
<td>.83</td>
<td>.79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCL-90-R</td>
<td>Derogatis, 1983</td>
<td>Work Problems</td>
<td>Combined genders</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------------</td>
<td>---------------</td>
<td>------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anxiety</td>
<td>.86</td>
<td>Anxiety</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Depression</td>
<td>.90</td>
<td>Depression</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hostility</td>
<td>.84</td>
<td>Hostility</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interpersonal Sensitivity</td>
<td>.84</td>
<td>Interpersonal Sensitivity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Obsessive-Compulsive</td>
<td>.86</td>
<td>Obsessive-Compulsive</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paranoid Ideation</td>
<td>.80</td>
<td>Paranoid Ideation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Psychoticism</td>
<td>.77</td>
<td>Psychoticism</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Somatization</td>
<td>.86</td>
<td>Somatization</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
status at the time of the evaluation was also recorded. Ten percent of the records were
coded by a second research assistant to assess interrater reliability. The mean Kappa
coefficient for dichotomous variables on the RRF was .77 and the mean intraclass
correlation coefficient for continuous variables on the RRF was .78.

In the civil forensic sample, criterion measures evaluating a number of
psychological and physical complaints were administered. Additionally, a number of
neuropsychological measures were administered to this sample; however, these will not
be used. A full reporting of these measures, along with internal consistency values and
descriptions, can be found in Table 4.

Data Analyses

Following the method used by Handel and colleagues (Handel et al., 2006),
increasing levels of random responses were randomly inserted at the individual protocol
level. By modifying protocols at the individual level it is highly unlikely that any two
protocols have an identical set of modified items. Using the original data sets, 30, 60, or
90% of the items for each individual protocol were randomly selected and randomly
changed to “true” or “false” responses, in order to simulate mild, moderate, and severe
random responding, respectively. Each ascending level of random responding was
inserted into a new copy of the original data set so that the same original protocol did not
have the exact same random responses when used to simulate varying degrees of random
responding.
<table>
<thead>
<tr>
<th>Criterion measure</th>
<th>Author</th>
<th>Construct</th>
<th>alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beck Depression Inventory (BDI)(^1)</td>
<td>Beck, et al., 1961</td>
<td>Symptoms of depression</td>
<td>.88</td>
</tr>
<tr>
<td>Detailed Assessment of Posttraumatic Stress (DAPS)</td>
<td>Briere, 2001</td>
<td>Substance Abuse</td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Suicidality</td>
<td>.94</td>
</tr>
<tr>
<td>Davidson Trauma Scale (DTS)(^3)</td>
<td>Davidson, 1996</td>
<td>Post-traumatic stress symptoms</td>
<td>.99</td>
</tr>
<tr>
<td>Memory Complaints Inventory (MCI)(^4)</td>
<td>Green, 2004</td>
<td>General Memory Problems</td>
<td>.88</td>
</tr>
<tr>
<td>Multifactor Health Inventory (MHI)(^4)</td>
<td>Hase, 1996</td>
<td>Anxiety Cluster</td>
<td>.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Depression Cluster</td>
<td>.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Memory Cluster</td>
<td>.65</td>
</tr>
<tr>
<td>Multidimensional Pain Inventory (MPI)(^5)</td>
<td>Kerns, Turk, &amp; Rudy, 1985</td>
<td>Affective Distress</td>
<td>.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td>General Activity Level</td>
<td>.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pain Interference</td>
<td>.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pain Severity</td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Perceived Life Control</td>
<td>.79</td>
</tr>
</tbody>
</table>
Examining the Attenuation of Validity Coefficients

The extra-test correlates and criterion measures described above were used to examine whether validity coefficients become attenuated as random responding increases. Extra-test criteria that correlate substantially with, and are conceptually related to, MMPI-2-RF scales were used in these analyses to examine the extent to which correlations are attenuated as a function of random responding. Correlations between MMPI-2-RF scale scores and criteria were compared and tested for significant differences across levels of inconsistent responding.

Examining the Effect of Random Responding on Substantive Scale Scores

Repeated measures analyses of variance, followed by planned comparisons were conducted to compare MMPI-2-RF T scores across different levels of random responding. Repeated Measure ANOVA was used as it reduces the error variance that is unrelated to the introduction of the manipulation, by focusing on the changes caused by the manipulation instead of individual differences (Field, 2005). Next, the percentage of individuals whose T-scores increase by 5 T-scores or more were compared across the levels of random responding followed by a comparison of the percentage of individuals producing clinically elevated scale scores across levels of random responding. To examine whether substantive scale score changes affect the pattern of scores on the H-O scales, Chi square analyses were also conducted. Specifically, the proportion of individuals for whom the highest scale on the protocol changes was compared across
levels of random responding using only participants who have at least one clinically elevated H-O scale, and for whom that score is at least 5 T-score points higher than the remaining H-O scales.

**Power Analyses**

Power analyses indicate that owing to the large sample sizes of the archival databases to be used in this study, sufficient power exists to obtain significant findings with small effect sizes if alpha is set at .05, an appropriate level given that internal replication was conducted with multiple samples.
Results

Does Prediction Bias Occur as a Function of Random Responding?

Conceptually-related variables were used to determine whether correlations between extra-test criteria and MMPI-2-RF Scales were affected by increasing levels of random response insertion in the five archival samples used in this study. Only criteria that were conceptually related to and had significant correlations (of .20 or higher) with the MMPI-2-RF hierarchical scales were chosen. Additionally, Williams’ (1959) t-test for dependent correlations was used across samples to test whether the correlations between MMPI-2-RF scales and criteria under various levels of random response insertion were statistically significantly different from the original correlation. In each sample, there were MMPI-2-RF scales that lacked appropriate criterion measures to examine them. These scales were omitted from the correlation tables without notations.

Correlations between MMPI-2-RF Scales and the relevant criteria for the college sample can be found in Table 5. In the table, we see that the first column is the MMPI-2-RF scale, then the criterion measure, followed by the correlation for that scale with the criterion at 0, 30, 60, and 90% random responses inserted. The asterisks show the results of Williams’ t-tests; with one asterisk denoting that the dependent correlation is different from the original correlation at the .05 level, two asterisks, that the correlations are different at the .01 level, and three asterisks denoting that the correlations are different at the .001 level. In the college sample, EID had a correlation of .70 with the Beck
Table 5. The Effects of 30%, 60% and 90% Random Response Insertion on MMPI-2 RF Scale Validity Coefficients in a College Sample

<table>
<thead>
<tr>
<th>Higher-Order Scales</th>
<th>0%</th>
<th>30%</th>
<th>60%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>EID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beck Depression Inventory</td>
<td>.70</td>
<td>.66***</td>
<td>.53***</td>
<td>.14***</td>
</tr>
<tr>
<td>STPI- Anxiety</td>
<td>.41</td>
<td>.41***</td>
<td>.35***</td>
<td>.10***</td>
</tr>
<tr>
<td>THD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Magical Ideation Scale</td>
<td>.58</td>
<td>.39***</td>
<td>.26***</td>
<td>.06***</td>
</tr>
<tr>
<td>BXD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIS-10 - General Impulsivity</td>
<td>.35</td>
<td>.30***</td>
<td>.22***</td>
<td>.07***</td>
</tr>
<tr>
<td>BIS-10 – Total Score</td>
<td>.41</td>
<td>.34***</td>
<td>.26***</td>
<td>.05***</td>
</tr>
<tr>
<td>Drug Abuse Screening Test</td>
<td>.50</td>
<td>.43***</td>
<td>.33***</td>
<td>.09***</td>
</tr>
<tr>
<td>Michigan Alcohol Screening Test</td>
<td>.41</td>
<td>.36***</td>
<td>.27***</td>
<td>.10***</td>
</tr>
<tr>
<td>Restructured Clinical Scales</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCd</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beck Depression Inventory</td>
<td>.70</td>
<td>.66***</td>
<td>.49***</td>
<td>.13***</td>
</tr>
<tr>
<td>STPI - Anxiety</td>
<td>.39</td>
<td>.38</td>
<td>.32***</td>
<td>.07***</td>
</tr>
<tr>
<td>RC1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSD</td>
<td>.63</td>
<td>.52***</td>
<td>.36***</td>
<td>.15***</td>
</tr>
<tr>
<td>RC2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beck Depression Inventory</td>
<td>.52</td>
<td>.40***</td>
<td>.25***</td>
<td>.05***</td>
</tr>
<tr>
<td>Internal State Scale - Depression</td>
<td>.39</td>
<td>.30***</td>
<td>.17***</td>
<td>.04***</td>
</tr>
<tr>
<td>RC3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Machiavellianism—IV</td>
<td>.54</td>
<td>.43***</td>
<td>.32***</td>
<td>.10***</td>
</tr>
<tr>
<td>RC4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIS-10 Total Score</td>
<td>.41</td>
<td>.32***</td>
<td>.25***</td>
<td>.06***</td>
</tr>
<tr>
<td>Drug Abuse Screening Test</td>
<td>.55</td>
<td>.49***</td>
<td>.35***</td>
<td>.06***</td>
</tr>
<tr>
<td>Family Functioning Scale</td>
<td>-.34</td>
<td>-.28***</td>
<td>-.22***</td>
<td>-.05***</td>
</tr>
<tr>
<td>Michigan Alcohol Screening Test</td>
<td>.38</td>
<td>.34**</td>
<td>.24***</td>
<td>.04***</td>
</tr>
<tr>
<td>Scale</td>
<td>Description</td>
<td>RC6</td>
<td>RC7</td>
<td>STPI - Anxiety</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------</td>
<td>-----</td>
<td>-------</td>
<td>---------------</td>
</tr>
<tr>
<td></td>
<td>Magical Ideation Scale</td>
<td>.44</td>
<td>.41</td>
<td>.32</td>
</tr>
<tr>
<td></td>
<td>Obsessive Compulsive Scale</td>
<td>.26***</td>
<td>.36***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>STPI - Anxiety</td>
<td>.17***</td>
<td>.30***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.03***</td>
<td>.08***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BIS-10 Motor Impulsivity</td>
<td>.47***</td>
<td>.40***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BIS-10 Total Scale</td>
<td>.19***</td>
<td>.17***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Internal State Scale - Hypomanic</td>
<td>.00***</td>
<td>.04***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Internal State Scale - Hypomanic</td>
<td>.05***</td>
<td>.07***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Beck Depression Inventory</td>
<td>.37</td>
<td>.37</td>
<td>.28</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.16***</td>
<td>.13***</td>
<td>.11***</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.06***</td>
<td>.05***</td>
<td>.06***</td>
</tr>
<tr>
<td></td>
<td>Beck Depression Inventory</td>
<td>.46</td>
<td>.46</td>
<td>.37</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.16***</td>
<td>.16***</td>
<td>.11***</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.07***</td>
<td>.07***</td>
<td>.07***</td>
</tr>
<tr>
<td></td>
<td>Beck Depression Inventory</td>
<td>.58</td>
<td>.58</td>
<td>.37</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.27***</td>
<td>.27***</td>
<td>.22***</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.09***</td>
<td>.09***</td>
<td>.09***</td>
</tr>
<tr>
<td></td>
<td>Beck Depression Inventory</td>
<td>.29</td>
<td>.29</td>
<td>.29</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.15***</td>
<td>.15***</td>
<td>.15***</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.07***</td>
<td>.07***</td>
<td>.07***</td>
</tr>
<tr>
<td></td>
<td>Beck Depression Inventory</td>
<td>.48</td>
<td>.48</td>
<td>.44</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.29***</td>
<td>.29***</td>
<td>.29***</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.05***</td>
<td>.05***</td>
<td>.05***</td>
</tr>
<tr>
<td></td>
<td>Beck Depression Inventory</td>
<td>.60</td>
<td>.60</td>
<td>.60</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.33***</td>
<td>.33***</td>
<td>.33***</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.09***</td>
<td>.09***</td>
<td>.09***</td>
</tr>
<tr>
<td></td>
<td>Beck Depression Inventory</td>
<td>.38</td>
<td>.38</td>
<td>.38</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.21***</td>
<td>.21***</td>
<td>.21***</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
</tr>
<tr>
<td></td>
<td>Beck Depression Inventory</td>
<td>.58</td>
<td>.58</td>
<td>.58</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.27***</td>
<td>.27***</td>
<td>.27***</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.09***</td>
<td>.09***</td>
<td>.09***</td>
</tr>
<tr>
<td></td>
<td>Beck Depression Inventory</td>
<td>.29</td>
<td>.29</td>
<td>.29</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.15***</td>
<td>.15***</td>
<td>.15***</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.07***</td>
<td>.07***</td>
<td>.07***</td>
</tr>
<tr>
<td></td>
<td>Beck Depression Inventory</td>
<td>.48</td>
<td>.48</td>
<td>.48</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.29***</td>
<td>.29***</td>
<td>.29***</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.05***</td>
<td>.05***</td>
<td>.05***</td>
</tr>
<tr>
<td></td>
<td>Beck Depression Inventory</td>
<td>.60</td>
<td>.60</td>
<td>.60</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.33***</td>
<td>.33***</td>
<td>.33***</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.09***</td>
<td>.09***</td>
<td>.09***</td>
</tr>
<tr>
<td></td>
<td>Beck Depression Inventory</td>
<td>.38</td>
<td>.38</td>
<td>.38</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.21***</td>
<td>.21***</td>
<td>.21***</td>
</tr>
<tr>
<td></td>
<td>Internal State Scale – Depressive</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
</tr>
</tbody>
</table>

Specific Problem Scales

<table>
<thead>
<tr>
<th>Scale</th>
<th>Description</th>
<th>MLSS</th>
<th>GICS</th>
<th>HPCS</th>
<th>SUIS</th>
<th>HLPIS</th>
<th>SFDIS</th>
<th>AXYS</th>
<th>ANPS</th>
<th>SUB</th>
<th>MASTE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLS</td>
<td>SSD</td>
<td>.47</td>
<td>.45</td>
<td>.56</td>
<td>.37</td>
<td>.46</td>
<td>.58</td>
<td>.29</td>
<td>.48</td>
<td>.60</td>
<td>.38</td>
</tr>
<tr>
<td>GIC</td>
<td>SSD</td>
<td>.35***</td>
<td>.30***</td>
<td>.41***</td>
<td>.16***</td>
<td>.35***</td>
<td>.48***</td>
<td>.20***</td>
<td>.40***</td>
<td>.49***</td>
<td>.29***</td>
</tr>
<tr>
<td>HPC</td>
<td>SSD</td>
<td>.19***</td>
<td>.17***</td>
<td>.24***</td>
<td>.13***</td>
<td>.16***</td>
<td>.27***</td>
<td>.15***</td>
<td>.29***</td>
<td>.33***</td>
<td>.21***</td>
</tr>
<tr>
<td>SUI</td>
<td>Beck Depression Inventory</td>
<td>.00***</td>
<td>.04***</td>
<td>.08***</td>
<td>.06***</td>
<td>.07***</td>
<td>.09***</td>
<td>.07***</td>
<td>.05***</td>
<td>.04***</td>
<td>.04***</td>
</tr>
<tr>
<td>HLP</td>
<td>Beck Depression Inventory</td>
<td>.05***</td>
<td>.07***</td>
<td>.07***</td>
<td>.05***</td>
<td>.09***</td>
<td>.09***</td>
<td>.07***</td>
<td>.05***</td>
<td>.04***</td>
<td>.04***</td>
</tr>
<tr>
<td>SFD</td>
<td>Beck Depression Inventory</td>
<td>.09***</td>
<td>.09***</td>
<td>.09***</td>
<td>.09***</td>
<td>.09***</td>
<td>.09***</td>
<td>.09***</td>
<td>.09***</td>
<td>.09***</td>
<td>.09***</td>
</tr>
<tr>
<td>AXY</td>
<td>STPI - Anxiety</td>
<td>.07***</td>
<td>.07***</td>
<td>.07***</td>
<td>.07***</td>
<td>.07***</td>
<td>.07***</td>
<td>.07***</td>
<td>.07***</td>
<td>.07***</td>
<td>.07***</td>
</tr>
<tr>
<td>ANP</td>
<td>STPI - Anger</td>
<td>.05***</td>
<td>.05***</td>
<td>.05***</td>
<td>.05***</td>
<td>.05***</td>
<td>.05***</td>
<td>.05***</td>
<td>.05***</td>
<td>.05***</td>
<td>.05***</td>
</tr>
<tr>
<td>SUB</td>
<td>Drug Abuse Screening Test</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
</tr>
<tr>
<td>MASTE</td>
<td>Michigan Alcohol Screening Test</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
<td>.04***</td>
</tr>
<tr>
<td>Scale</td>
<td>STPI - Anger</td>
<td>Internal State Scale - Hypomanic</td>
<td>BIS/BAS Fun Seeking</td>
<td>Family Functioning Scale</td>
<td>BIS/BAS - Inhibition</td>
<td>Fears Questionnaire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------------</td>
<td>---------------------------------</td>
<td>---------------------</td>
<td>--------------------------</td>
<td>---------------------</td>
<td>--------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGG</td>
<td>.39</td>
<td>.33***</td>
<td>.19***</td>
<td>.10***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACT</td>
<td>.30</td>
<td>.23***</td>
<td>.14***</td>
<td>.01***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FML</td>
<td>.25</td>
<td>.18***</td>
<td>.12***</td>
<td>-.02***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHY</td>
<td>.37</td>
<td>.28***</td>
<td>.21***</td>
<td>.09***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BIS/BAS - Behavioral Inhibition/Activation System; BIS-10 - Barratt Impulsivity Scale—10; SSD - Screener for Somatoform Disorders; STPI – State Trait Personality Inventory

Statistically significantly different from 0% random response condition:
* - p < .05 level; ** - p < .01, *** - p < .001
Depression Inventory. The correlation showed a statistically significant, though not large drop at 30% random response insertion (to .66), a modest drop at 60% (to .53), and was substantially attenuated at 90% random response insertion (r = .14). Generally, most correlations between the MMPI-2-RF Scales and relevant criterion measures showed small but statistically significant differences even at 30%; however, on average, the degree of attenuation of the correlations was not large, with one correlation changing as little as .01 (e.g., RCd and STPI-Anxiety). There were some notable exceptions (e.g., the correlation between THD and the Magical Ideation Scale dropped from .58 to .39 with only 30% random responses inserted). At 60% random response insertion, all of the correlations between MMPI-2-RF Scales and criterion measures showed modest attenuation, and all correlations were statistically significantly lower than the original correlation. At 90% random response insertion, all correlations were significantly attenuated, with only three correlations above .10 (i.e., EID and BDI (.14), RCd and BDI (.13), and RC 9 with BIS-10 Motor Impulsivity (.11)), the conventional cutoff for a small correlation (Cohen, 1988). In addition, there were no observable differences in the pattern of changes in the correlations regardless of the length of the scales; overall, the correlations for the Higher Order scales did not appear to be more robust than those for the Specific Problem scales, or vice versa. On the other hand, scales composed of items with lower rates of endorsement (THD, RC6, RC8, SUI) showed more rapid attenuation of correlations than other scales did.

Outpatient correlations between the MMPI-2-RF Scales and the relevant criteria from the PDF Scales can be found in Table 6. In this sample, many correlations were
quite robust at 30% random response insertion. However, out of the 50 correlations between extra-test criteria and MMPI-2-RF Scales that were examined, only nine correlations were not statistically different from the original correlation at 30% random response insertion. At 60% random response insertion, all of the correlations between MMPI-2-RF Scales and criterion measures showed at least modest attenuation, and all but five correlations were highly statistically significantly lower than the original correlation. At 90% random response insertion, all correlations were statistically significant and dramatically lowered. Although all of the criteria were chosen in part because they had correlations greater than .20, when 90% of valid responses were replaced with random ones, only one correlation was greater than .20 (i.e., EID was correlated with SCL-R-90 Depression at .23 at 90% random response insertion, though the original correlation was .79). However, 18 out of 50 correlations remained above .10 in the 90% random response insertion condition. The length of the scales did not appear to affect the pattern or rapidity of correlation attenuation.

For the inpatient sample, correlations between the MMPI-2-RF Scales and the relevant conceptually-related criteria can be found in Table 7. As with the previous two samples, many of the correlations were quite robust at 30% random response insertion. In fact, at 30% random response insertion, eight of out of the 18 correlations between extra-test criteria and MMPI-2-RF Scales were not statistically significantly different from the original correlation, and one was only statistically significantly different at the $p < .05$ level. However, at 60% random response insertion, all but one of the correlations
Table 6. The Effects of 30%, 60% and 90% Random Response Insertion on MMPI-2 RF Scale Validity Coefficients in an Outpatient Sample

<table>
<thead>
<tr>
<th>Scale Description</th>
<th>0%</th>
<th>30%</th>
<th>60%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher Order Scales</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EID PDF Anxious</td>
<td>.35</td>
<td>.34</td>
<td>.25***</td>
<td>.12***</td>
</tr>
<tr>
<td>PDF Depressed</td>
<td>.49</td>
<td>.47*</td>
<td>.36***</td>
<td>.15***</td>
</tr>
<tr>
<td>SCL-R-90 Obsessive-Compulsive</td>
<td>.65</td>
<td>.63**</td>
<td>.53***</td>
<td>.18***</td>
</tr>
<tr>
<td>SCL-R-90 Depression</td>
<td>.79</td>
<td>.75***</td>
<td>.63***</td>
<td>.23***</td>
</tr>
<tr>
<td>SCL-R-90 Anxiety</td>
<td>.64</td>
<td>.60***</td>
<td>.54***</td>
<td>.19***</td>
</tr>
<tr>
<td>THD SCL-R-90 Paranoid Ideation</td>
<td>.52</td>
<td>.37***</td>
<td>.28***</td>
<td>.04***</td>
</tr>
<tr>
<td>SCL-R-90 Psychoticism</td>
<td>.46</td>
<td>.35***</td>
<td>.24***</td>
<td>.04***</td>
</tr>
<tr>
<td>Restructured Clinical Scales</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BXD PDF Aggressive</td>
<td>.37</td>
<td>.36</td>
<td>.25***</td>
<td>.03***</td>
</tr>
<tr>
<td>PDF Antisocial</td>
<td>.36</td>
<td>.35</td>
<td>.25***</td>
<td>.04***</td>
</tr>
<tr>
<td>RCd PDF Depressed</td>
<td>.46</td>
<td>.43</td>
<td>.36***</td>
<td>.10***</td>
</tr>
<tr>
<td>SCL-R-90 Depression</td>
<td>.81</td>
<td>.76***</td>
<td>.65***</td>
<td>.20***</td>
</tr>
<tr>
<td></td>
<td>RC1</td>
<td>RC2</td>
<td>RC4</td>
<td>RC6</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>SCL-R-90 Anxiety</td>
<td>.63</td>
<td>.46</td>
<td>.45</td>
<td>.64</td>
</tr>
<tr>
<td></td>
<td>.59*</td>
<td>.44</td>
<td>.43</td>
<td>.58***</td>
</tr>
<tr>
<td></td>
<td>.51***</td>
<td>.35***</td>
<td>.27***</td>
<td>.42***</td>
</tr>
<tr>
<td></td>
<td>.19***</td>
<td>.13***</td>
<td>.11***</td>
<td>.12***</td>
</tr>
<tr>
<td>SCL-R-90 Depression</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCL-R-90 Phobic Anxiety</td>
<td>.52</td>
<td>.58</td>
<td>.60</td>
<td>.52</td>
</tr>
<tr>
<td></td>
<td>.52***</td>
<td>.52***</td>
<td>.54***</td>
<td>.54***</td>
</tr>
<tr>
<td></td>
<td>.41***</td>
<td>.41***</td>
<td>.43***</td>
<td>.43***</td>
</tr>
<tr>
<td></td>
<td>.14***</td>
<td>.14***</td>
<td>.12***</td>
<td>.12***</td>
</tr>
<tr>
<td>SCL-R-90 Obsessive-Compulsive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCL-R-90 Anxiety</td>
<td>.60</td>
<td>.52</td>
<td>.52</td>
<td>.52</td>
</tr>
<tr>
<td></td>
<td>.54***</td>
<td>.52***</td>
<td>.54***</td>
<td>.52***</td>
</tr>
<tr>
<td></td>
<td>.43***</td>
<td>.43***</td>
<td>.43***</td>
<td>.43***</td>
</tr>
<tr>
<td></td>
<td>.12***</td>
<td>.12***</td>
<td>.12***</td>
<td>.12***</td>
</tr>
<tr>
<td>SCL-R-90 Phobic Anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific Problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDF Somatic Symptoms</td>
<td>.45</td>
<td>.45</td>
<td>.46</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>.41*</td>
<td>.41*</td>
<td>.46*</td>
<td>.41*</td>
</tr>
<tr>
<td></td>
<td>.29***</td>
<td>.29***</td>
<td>.40***</td>
<td>.29***</td>
</tr>
<tr>
<td>Scale</td>
<td>Test</td>
<td>r</td>
<td>t</td>
<td>p</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------------------------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>SCL-R-90 Somatization</td>
<td>GIC PDF Somatic Symptoms</td>
<td>.65</td>
<td>.57***</td>
<td>.39***</td>
</tr>
<tr>
<td></td>
<td>SCL-R-90 Somatization</td>
<td>.58</td>
<td>.43***</td>
<td>.29***</td>
</tr>
<tr>
<td></td>
<td>HPC PDF Somatic Symptoms</td>
<td>.35</td>
<td>.27***</td>
<td>.16***</td>
</tr>
<tr>
<td></td>
<td>SCL-R-90 Somatization</td>
<td>.66</td>
<td>.55***</td>
<td>.34***</td>
</tr>
<tr>
<td></td>
<td>SUI PDF Depressed</td>
<td>.35</td>
<td>.15***</td>
<td>.16***</td>
</tr>
<tr>
<td></td>
<td>SCL-R-90 Depression</td>
<td>.48</td>
<td>.33**</td>
<td>.20***</td>
</tr>
<tr>
<td></td>
<td>HLP PDF Pessimistic</td>
<td>.26</td>
<td>.24</td>
<td>.09***</td>
</tr>
<tr>
<td></td>
<td>PDF Depressed</td>
<td>.36</td>
<td>.34***</td>
<td>.13***</td>
</tr>
<tr>
<td></td>
<td>SCL-R-90 Depression</td>
<td>.57</td>
<td>.44***</td>
<td>.29***</td>
</tr>
<tr>
<td></td>
<td>SFD PDF Insecure</td>
<td>.32</td>
<td>.25***</td>
<td>.17***</td>
</tr>
<tr>
<td></td>
<td>STW PDF Anxious</td>
<td>.30</td>
<td>.24***</td>
<td>.09***</td>
</tr>
<tr>
<td></td>
<td>AXY PDF Anxious</td>
<td>.31</td>
<td>.23***</td>
<td>.17***</td>
</tr>
<tr>
<td></td>
<td>SCL-R-90 Anxiety</td>
<td>.66</td>
<td>.50***</td>
<td>.32***</td>
</tr>
<tr>
<td></td>
<td>ANP SCL-R-90 Hostility</td>
<td>.62</td>
<td>.54*</td>
<td>.33***</td>
</tr>
<tr>
<td>Scale</td>
<td>Description</td>
<td>M1</td>
<td>M2</td>
<td>M3</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>BRF</td>
<td>SCL-R-90 Phobic Anxiety</td>
<td>.50</td>
<td>.32***</td>
<td>.15***</td>
</tr>
<tr>
<td>MSF</td>
<td>SCL-R-90 Phobic Anxiety</td>
<td>.32</td>
<td>.26**</td>
<td>.17***</td>
</tr>
<tr>
<td>JCP</td>
<td>PDF Aggressive</td>
<td>.31</td>
<td>.28**</td>
<td>.17***</td>
</tr>
<tr>
<td></td>
<td>PDF Antisocial</td>
<td>.33</td>
<td>.30***</td>
<td>.20***</td>
</tr>
<tr>
<td>AGG</td>
<td>PDF Angry Resentment</td>
<td>.23</td>
<td>.22**</td>
<td>.13**</td>
</tr>
<tr>
<td></td>
<td>PDF Aggressive</td>
<td>.28</td>
<td>.23***</td>
<td>.19**</td>
</tr>
<tr>
<td></td>
<td>SCL-R-90 Hostility</td>
<td>.64</td>
<td>.52***</td>
<td>.32***</td>
</tr>
<tr>
<td>FML</td>
<td>PDF Family Problems</td>
<td>.32</td>
<td>.24***</td>
<td>.15***</td>
</tr>
<tr>
<td>IPP</td>
<td>PDF Passive-Submissive</td>
<td>.31</td>
<td>.27</td>
<td>.25*</td>
</tr>
<tr>
<td>SAV</td>
<td>PDF Introverted</td>
<td>.29</td>
<td>.24***</td>
<td>.24***</td>
</tr>
<tr>
<td>SHY</td>
<td>SCL-R-90 Interpersonal Sensitivity</td>
<td>.50</td>
<td>.41***</td>
<td>.29***</td>
</tr>
</tbody>
</table>

PDF – Patient Description Form Scales
Statistically significantly different from 0% random response condition:
* - p < .05 level; ** - p < .01, *** - p < .001
Table 7. The Effects of 30, 60, and 90% Random Response Insertion on MMPI-2 RF Scale Validity Coefficients in an Inpatient Sample

<table>
<thead>
<tr>
<th></th>
<th>0%</th>
<th>30%</th>
<th>60%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Higher Order</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>.30</td>
<td>.29</td>
<td>.22***</td>
<td>.09***</td>
</tr>
<tr>
<td>Suicidality</td>
<td>.31</td>
<td>.30</td>
<td>.25***</td>
<td>.12***</td>
</tr>
<tr>
<td>THD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hallucinations</td>
<td>.28</td>
<td>.24</td>
<td>.14***</td>
<td>.04***</td>
</tr>
<tr>
<td>Delusions</td>
<td>.23</td>
<td>.22</td>
<td>.10***</td>
<td>.01***</td>
</tr>
<tr>
<td>BXD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Substance Abuse</td>
<td>.51</td>
<td>.45*</td>
<td>.35***</td>
<td>.12***</td>
</tr>
<tr>
<td>Criminal Justice Involvement</td>
<td>.23</td>
<td>.22</td>
<td>.16***</td>
<td>.07***</td>
</tr>
<tr>
<td><strong>Restructured Clinical Scales</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCd</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>.32</td>
<td>.30**</td>
<td>.25***</td>
<td>.08***</td>
</tr>
<tr>
<td>RC2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>.26</td>
<td>.23**</td>
<td>.14***</td>
<td>.05***</td>
</tr>
<tr>
<td>RC4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Substance Abuse</td>
<td>.55</td>
<td>.50***</td>
<td>.39***</td>
<td>.14***</td>
</tr>
<tr>
<td>Criminal Justice Involvement</td>
<td>.22</td>
<td>.21</td>
<td>.14***</td>
<td>.05***</td>
</tr>
<tr>
<td>RC6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delusions</td>
<td>.25</td>
<td>.22</td>
<td>.12**</td>
<td>&lt;.01***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>--------------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>RC7</td>
<td>Intrusive Ideation</td>
<td>.21</td>
<td>.20</td>
<td>.16***</td>
</tr>
<tr>
<td>RC8</td>
<td>Hallucinations</td>
<td>.26</td>
<td>.22**</td>
<td>.13***</td>
</tr>
</tbody>
</table>

Specific Problems

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HPC</td>
<td>Pain</td>
<td>.26</td>
<td>.20***</td>
<td>.14***</td>
<td>.03***</td>
</tr>
<tr>
<td>SUI</td>
<td>Suicidality</td>
<td>.55</td>
<td>.42***</td>
<td>.26***</td>
<td>.09***</td>
</tr>
<tr>
<td>SFD</td>
<td>Depression</td>
<td>.31</td>
<td>.24***</td>
<td>.16***</td>
<td>.05***</td>
</tr>
<tr>
<td>AXY</td>
<td>Intrusive Ideation</td>
<td>.28</td>
<td>.23***</td>
<td>.15***</td>
<td>.01***</td>
</tr>
<tr>
<td>SUB</td>
<td>Substance Abuse</td>
<td>.59</td>
<td>.50***</td>
<td>.38***</td>
<td>.08***</td>
</tr>
</tbody>
</table>

* All correlations are with Record Review Form Scales
Statistically significantly different from 0% random response condition:
* - p < .05 level; ** - p < .01, *** - p < .001
between MMPI-2-RF Scales and criterion measures were statistically significant (at \( p < .001 \)) and at 90% random response insertion, all correlations were greatly attenuated and statistically significantly different from the original correlation. In the inpatient sample, no correlations remained above .10 when 90% of responses were replaced with random responses. In addition, similarly to the first two samples, there was no observable difference how correlations attenuated relative to scale length.

Correlations between conceptually-related criteria can be found in Table 8 for the criminal forensic sample. At 30% random response insertion, six of out of the 21 examined correlations were not statistically significantly different from the original correlation, and four correlations were only statistically significantly different at the \( p < .05 \) level. Yet in the 60% random response insertion, only three of the correlations between MMPI-2-RF Scales and criterion measures were not highly statistically significantly different from their original counterparts. At 90% random response insertion, all correlations were statistically and clinically significantly attenuated. For this sample, there were four correlations that were still above .10 in the 90% random response insertion condition. As with the other samples, there was no observable difference in the patterns of attenuation in the correlations regardless of the length of the scales.

Correlations between the MMPI-2-RF Scales and extra-test criteria for the civil forensic sample can be found in Table 9. In this sample, 31 out of the 37 correlations were statistically significantly different from the original correlation at 30% random response insertion, although five of the correlations were only statistically significantly
Table 8. The Effects of 30%, 60% and 90% Random Response Insertion on MMPI-2 RF Scale Validity Coefficients in a Criminal Forensic Sample

<table>
<thead>
<tr>
<th></th>
<th>0%</th>
<th>30%</th>
<th>60%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Higher Order Scales</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRRF Dysphoria</td>
<td>.25</td>
<td>.25</td>
<td>.20*</td>
<td>.11***</td>
</tr>
<tr>
<td>BXD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRRF Juvenile</td>
<td>.43</td>
<td>.40*</td>
<td>.30***</td>
<td>.05***</td>
</tr>
<tr>
<td>Delinquency/Conduct</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRRF Substance Abuse</td>
<td>.47</td>
<td>.42***</td>
<td>.28***</td>
<td>.08***</td>
</tr>
<tr>
<td>FRRF Drug Dependency</td>
<td>.40</td>
<td>.36**</td>
<td>.21***</td>
<td>.11***</td>
</tr>
<tr>
<td>FRRF Criminal History</td>
<td>.23</td>
<td>.22</td>
<td>.14***</td>
<td>.05***</td>
</tr>
<tr>
<td><strong>Restructured Clinical Scales</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCd</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRRF Dysphoria</td>
<td>.25</td>
<td>.22**</td>
<td>.20*</td>
<td>.04***</td>
</tr>
<tr>
<td>RC1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRRF Medical Problems</td>
<td>.27</td>
<td>.24</td>
<td>.14***</td>
<td>.01***</td>
</tr>
<tr>
<td>RC2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRRF Dysphoria</td>
<td>.28</td>
<td>.23**</td>
<td>.14***</td>
<td>.06***</td>
</tr>
<tr>
<td>RC4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRRF Juvenile</td>
<td>.43</td>
<td>.39**</td>
<td>.31***</td>
<td>.06***</td>
</tr>
<tr>
<td>Delinquency/Conduct</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRRF Substance Abuse</td>
<td>.51</td>
<td>.46***</td>
<td>.35***</td>
<td>.10***</td>
</tr>
<tr>
<td>FRRF Drug Dependency</td>
<td>.43</td>
<td>.38***</td>
<td>.27***</td>
<td>.15***</td>
</tr>
<tr>
<td>FRRF Criminal History</td>
<td>.26</td>
<td>.23*</td>
<td>.16***</td>
<td>.03***</td>
</tr>
<tr>
<td><strong>Specific Problems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRRF Medical Problems</td>
<td>.27</td>
<td>.24</td>
<td>.19**</td>
<td>.03***</td>
</tr>
<tr>
<td>GIC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRRF Medical Problems</td>
<td>.22</td>
<td>.13***</td>
<td>.07***</td>
<td>-.04***</td>
</tr>
<tr>
<td>FRRF</td>
<td>Medical Problems</td>
<td>.22</td>
<td>.17*</td>
<td>.10***</td>
</tr>
<tr>
<td>------</td>
<td>------------------</td>
<td>-----</td>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td>NUC</td>
<td>FRRF Medical Problems</td>
<td>.24</td>
<td>.23</td>
<td>.09***</td>
</tr>
<tr>
<td>JCP</td>
<td>FRRF Juvenile Delinquency/Conduct</td>
<td>.54</td>
<td>.46***</td>
<td>.31***</td>
</tr>
<tr>
<td>SUB</td>
<td>FRRF Substance Abuse</td>
<td>.55</td>
<td>.44***</td>
<td>.31***</td>
</tr>
<tr>
<td></td>
<td>FRRF Drug Dependency</td>
<td>.55</td>
<td>.46***</td>
<td>.33***</td>
</tr>
<tr>
<td>AGG</td>
<td>FRRF Violent/Abusive</td>
<td>.20</td>
<td>.17</td>
<td>.09***</td>
</tr>
<tr>
<td>FML</td>
<td>FRRF Abused</td>
<td>.28</td>
<td>.24*</td>
<td>.16***</td>
</tr>
</tbody>
</table>

FRRF – Forensic Record Review Form Scales; Statistically significantly different from 0% random response condition:* - $p < .05$ level; ** - $p < .01$, *** - $p < .001$
Table 9. The Effects of 30%, 60% and 90% Random Response Insertion on MMPI-2 RF Scale Validity Coefficients in an Civil Forensic Sample

<table>
<thead>
<tr>
<th>Higher Order Scales</th>
<th>EID</th>
<th>0%</th>
<th>30%</th>
<th>60%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beck Depression Inventory</td>
<td>0.62</td>
<td>0.62</td>
<td>0.52***</td>
<td>0.22***</td>
<td></td>
</tr>
<tr>
<td>DAPS Suicidality</td>
<td>0.35</td>
<td>0.36</td>
<td>0.30***</td>
<td>0.12***</td>
<td></td>
</tr>
<tr>
<td>MHI Anxiety</td>
<td>0.66</td>
<td>0.63***</td>
<td>0.54***</td>
<td>0.18***</td>
<td></td>
</tr>
<tr>
<td>MHI Depression</td>
<td>0.69</td>
<td>0.66***</td>
<td>0.56***</td>
<td>0.23***</td>
<td></td>
</tr>
<tr>
<td>Davidson Trauma Scale</td>
<td>0.58</td>
<td>0.58</td>
<td>0.45***</td>
<td>0.18***</td>
<td></td>
</tr>
<tr>
<td>MPI Affective Distress</td>
<td>0.50</td>
<td>0.50</td>
<td>0.42***</td>
<td>0.16***</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Restructured Clinical Scales</th>
<th>BXD</th>
<th>0%</th>
<th>30%</th>
<th>60%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAPS Substance Abuse</td>
<td>0.34</td>
<td>0.29***</td>
<td>0.22***</td>
<td>0.11***</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Revised Scales</th>
<th>RDd</th>
<th>0%</th>
<th>30%</th>
<th>60%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beck Depression Inventory</td>
<td>0.64</td>
<td>0.62**</td>
<td>0.50***</td>
<td>0.14***</td>
<td></td>
</tr>
<tr>
<td>MHI Anxiety</td>
<td>0.65</td>
<td>0.60***</td>
<td>0.50***</td>
<td>0.13***</td>
<td></td>
</tr>
<tr>
<td>MHI Depression</td>
<td>0.70</td>
<td>0.65***</td>
<td>0.52***</td>
<td>0.14***</td>
<td></td>
</tr>
<tr>
<td>Davidson Trauma Scale</td>
<td>0.59</td>
<td>0.56***</td>
<td>0.40***</td>
<td>0.18***</td>
<td></td>
</tr>
<tr>
<td>MPI Affective Distress</td>
<td>0.51</td>
<td>0.48***</td>
<td>0.37***</td>
<td>0.17***</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Revised Scales</th>
<th>RC1</th>
<th>0%</th>
<th>30%</th>
<th>60%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHI Memory</td>
<td>0.40</td>
<td>0.37**</td>
<td>0.25***</td>
<td>0.14***</td>
<td></td>
</tr>
<tr>
<td>MCI Global Memory Problems</td>
<td>0.42</td>
<td>0.37***</td>
<td>0.28***</td>
<td>0.15***</td>
<td></td>
</tr>
<tr>
<td>MPI Pain Severity</td>
<td>0.30</td>
<td>0.26***</td>
<td>0.20***</td>
<td>0.10***</td>
<td></td>
</tr>
<tr>
<td>MPI Pain Interference</td>
<td>0.30</td>
<td>0.27**</td>
<td>0.23***</td>
<td>0.09***</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Revised Scales</th>
<th>RC2</th>
<th>0%</th>
<th>30%</th>
<th>60%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beck Depression Inventory</td>
<td>0.48</td>
<td>0.44***</td>
<td>0.30***</td>
<td>0.06***</td>
<td></td>
</tr>
<tr>
<td>Scale</td>
<td>MLS</td>
<td>CIC</td>
<td>HPC</td>
<td>NUC</td>
<td>COG</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>DAPS Suicidality</td>
<td>.28</td>
<td>.27</td>
<td>.19***</td>
<td>.06***</td>
<td>.28</td>
</tr>
<tr>
<td>MHI Depression</td>
<td>.48</td>
<td>.46</td>
<td>.34***</td>
<td>.10***</td>
<td>.48</td>
</tr>
<tr>
<td>MPI General Activity Level</td>
<td>-.30</td>
<td>-.27*</td>
<td>-.20***</td>
<td>-.06***</td>
<td>-.30</td>
</tr>
<tr>
<td>RC4 DAPS Substance Abuse</td>
<td>.36</td>
<td>.32**</td>
<td>.19***</td>
<td>.07***</td>
<td>.36</td>
</tr>
<tr>
<td>RC7 MHI Anxiety</td>
<td>.66</td>
<td>.61***</td>
<td>.47***</td>
<td>.17***</td>
<td>.66</td>
</tr>
<tr>
<td>Davidson Trauma scale</td>
<td>.52</td>
<td>.50*</td>
<td>.39***</td>
<td>.14***</td>
<td>.52</td>
</tr>
<tr>
<td>MPI Affective Distress</td>
<td>.44</td>
<td>.42***</td>
<td>.32***</td>
<td>.11***</td>
<td>.44</td>
</tr>
<tr>
<td>Specific Problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLS MPI Pain Severity</td>
<td>.33</td>
<td>.26***</td>
<td>.16***</td>
<td>.06***</td>
<td>.33</td>
</tr>
<tr>
<td>MPI Interference</td>
<td>.41</td>
<td>.31***</td>
<td>.21***</td>
<td>.06***</td>
<td>.41</td>
</tr>
<tr>
<td>MPI Perceived Life Control</td>
<td>-.32</td>
<td>-.25*</td>
<td>-.18***</td>
<td>-.10***</td>
<td>-.32</td>
</tr>
<tr>
<td>MPI General Activity Level</td>
<td>-.27</td>
<td>-.23**</td>
<td>-.15***</td>
<td>-.04***</td>
<td>-.27</td>
</tr>
<tr>
<td>GIC MPI Pain Severity</td>
<td>.20</td>
<td>.16*</td>
<td>.08***</td>
<td>.01***</td>
<td>.20</td>
</tr>
<tr>
<td>MPI Interference</td>
<td>.21</td>
<td>.17*</td>
<td>.07***</td>
<td>.01***</td>
<td>.21</td>
</tr>
<tr>
<td>HPC MPI Pain Severity</td>
<td>.20</td>
<td>.15***</td>
<td>.08***</td>
<td>.06***</td>
<td>.20</td>
</tr>
<tr>
<td>NUC MHI Memory</td>
<td>.37</td>
<td>.31***</td>
<td>.20***</td>
<td>.06***</td>
<td>.37</td>
</tr>
<tr>
<td>MCI Global Memory Problems</td>
<td>.37</td>
<td>.31***</td>
<td>.21***</td>
<td>.08***</td>
<td>.37</td>
</tr>
<tr>
<td>MPI Pain Severity</td>
<td>.28</td>
<td>.21***</td>
<td>.20***</td>
<td>.08***</td>
<td>.28</td>
</tr>
<tr>
<td>COG MHI Memory</td>
<td>.61</td>
<td>.53***</td>
<td>.35***</td>
<td>.08***</td>
<td>.61</td>
</tr>
<tr>
<td>MHI Global Memory Problems</td>
<td>.58</td>
<td>.50***</td>
<td>.33***</td>
<td>.10***</td>
<td>.58</td>
</tr>
<tr>
<td>DAPS Impairment</td>
<td>.31</td>
<td>.27**</td>
<td>.16***</td>
<td>.04***</td>
<td>.31</td>
</tr>
<tr>
<td></td>
<td>DAPS Suicidality</td>
<td>.74</td>
<td>.46***</td>
<td>.24***</td>
<td>.04***</td>
</tr>
<tr>
<td>-------</td>
<td>------------------</td>
<td>-----</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td>0%</td>
<td>30%</td>
<td>60%</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>HLP</td>
<td>Beck Depression Inventory</td>
<td>.50</td>
<td>.40***</td>
<td>.23***</td>
<td>.10***</td>
</tr>
<tr>
<td></td>
<td>MHI Depression</td>
<td>.52</td>
<td>.41***</td>
<td>.28***</td>
<td>.11***</td>
</tr>
<tr>
<td></td>
<td>MPI Perceived Life Control</td>
<td>-.41</td>
<td>-.32***</td>
<td>-.20***</td>
<td>-.10***</td>
</tr>
<tr>
<td>SFD</td>
<td>Beck Depression Inventory</td>
<td>.55</td>
<td>.45***</td>
<td>.31***</td>
<td>.08***</td>
</tr>
<tr>
<td></td>
<td>MHI Anxiety</td>
<td>.50</td>
<td>.41***</td>
<td>.27***</td>
<td>.10***</td>
</tr>
<tr>
<td></td>
<td>MHI Depression</td>
<td>.56</td>
<td>.48***</td>
<td>.28***</td>
<td>.08***</td>
</tr>
<tr>
<td>NFC</td>
<td>MHI Anxiety</td>
<td>.49</td>
<td>.45***</td>
<td>.33***</td>
<td>.10***</td>
</tr>
<tr>
<td>STW</td>
<td>MHI Anxiety</td>
<td>.57</td>
<td>.49***</td>
<td>.30***</td>
<td>.05***</td>
</tr>
<tr>
<td>AXY</td>
<td>MHI Anxiety</td>
<td>.67</td>
<td>.54***</td>
<td>.31***</td>
<td>.08***</td>
</tr>
<tr>
<td></td>
<td>MPI Affective Distress</td>
<td>.42</td>
<td>.34***</td>
<td>.19***</td>
<td>.04***</td>
</tr>
<tr>
<td>SUB</td>
<td>DAPS Substance Abuse</td>
<td>.54</td>
<td>.39***</td>
<td>.26***</td>
<td>.07***</td>
</tr>
</tbody>
</table>

DAPS – Detailed Assessment of Post-traumatic Stress; MCI – Memory Complaints Inventory; MHI – Multifactor Health Inventory; MPI – Multidimensional Pain Inventory
Statistically significantly different from 0% random response condition:
* - p < .05 level; ** - p < .01, *** - p < .001
different at the $p < .05$ level. All of the correlations were highly statistically significant at 60% and 90% random response insertion. For this sample, four correlations were greater than .10 in the 90% random response insertion condition. The length of the scales did not appear to affect the patterns of attenuation in the correlations.

**Does Random Responding Distort Scores on Substantive Scales in a Manner that would Affect Their Interpretation?**

**Individual scale analyses.**

Repeated measures ANOVAs were used examine how individual scale scores are distorted under greater amounts of random responding. These analyses, along with means and standard deviations on the MMPI-2-RF Scales, the results of post-hoc comparisons and associated effect sizes are reported for the college sample in Table 10. Looking across Table 10, there were significant differences in mean scale scores at various levels of random response insertion on EID ($F(1.66,2478) = 1296.04$, $p < .001$). The results of the post-hoc tests for EID were all significant at $p < .001$. These contrasts illustrate the degree to which mean scores on EID became increasingly inflated as greater proportions of random responses were inserted. Means scores were significantly higher in the 30% random response insertion condition (55.55 [8.2]; $t = 32.57$, $df = 1490$, $p<.017$, $d=.38$) than when no random responses had been inserted (51.11 [11.6]). The scores became more discrepant from the unaltered scores in the 60% random response condition (59.33 [5.7]; $t = 38.51$, $df = 1490$, $p<.001$, $d=.71$), and highly discrepant scores were found in the 90% random response condition (63.34 [4.4]; $t = 41.19$, $df = 1490$, $p<.001$, $d=.79$).
Table 10. The Effects of 30%, 60%, and 90% Random Response Insertion on MMPI-2-RF Scale Scores in a College Sample

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>$\eta^2_{PARTIAL}$</th>
<th>None Mean(SD)</th>
<th>30% Mean(SD)</th>
<th>60% Mean(SD)</th>
<th>90% Mean(SD)</th>
<th>d</th>
<th>30% Mean(SD)</th>
<th>60% Mean(SD)</th>
<th>90% Mean(SD)</th>
<th>d</th>
<th>30% Mean(SD)</th>
<th>60% Mean(SD)</th>
<th>90% Mean(SD)</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>EID$^a$</td>
<td>1296.04</td>
<td>1.66,2478</td>
<td>&lt;.001</td>
<td>.47</td>
<td>51.11(11.6)</td>
<td>55.55(8.2)***</td>
<td>59.33(5.7)***</td>
<td>63.34(4.4)***</td>
<td>.71</td>
<td>51.11(11.6)</td>
<td>55.55(8.2)***</td>
<td>59.33(5.7)***</td>
<td>63.34(4.4)***</td>
<td>.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>THD$^b$</td>
<td>5614.88</td>
<td>2.61,3892</td>
<td>&lt;.001</td>
<td>.79</td>
<td>52.78(10.7)</td>
<td>65.72(8.6)***</td>
<td>76.45(8.6)***</td>
<td>87.38(7.8)***</td>
<td>2.21</td>
<td>65.72(8.6)***</td>
<td>76.45(8.6)***</td>
<td>87.38(7.8)***</td>
<td>2.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BXD$^a$</td>
<td>620.25</td>
<td>2.16,3224</td>
<td>&lt;.001</td>
<td>.29</td>
<td>53.88(10.3)</td>
<td>56.82(8.5)***</td>
<td>59.83(6.9)***</td>
<td>62.96(6.2)***</td>
<td>.88</td>
<td>56.82(8.5)***</td>
<td>59.83(6.9)***</td>
<td>62.96(6.2)***</td>
<td>.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC$^d$</td>
<td>656.17</td>
<td>1.71,2546</td>
<td>&lt;.001</td>
<td>.31</td>
<td>54.68(11.1)</td>
<td>57.98(7.6)***</td>
<td>60.40(5.5)***</td>
<td>63.23(4.5)***</td>
<td>.77</td>
<td>57.98(7.6)***</td>
<td>60.40(5.5)***</td>
<td>63.23(4.5)***</td>
<td>.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC$^a$</td>
<td>2807.38</td>
<td>2.21,3297</td>
<td>&lt;.001</td>
<td>.65</td>
<td>53.77(11.1)</td>
<td>61.36(8.0)***</td>
<td>67.63(6.5)***</td>
<td>73.70(5.9)***</td>
<td>1.80</td>
<td>61.36(8.0)***</td>
<td>67.63(6.5)***</td>
<td>73.70(5.9)***</td>
<td>1.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC$^b$</td>
<td>1517.66</td>
<td>2.42,3598</td>
<td>&lt;.001</td>
<td>.51</td>
<td>48.36(10.5)</td>
<td>54.07(9.2)***</td>
<td>59.71(8.2)***</td>
<td>65.12(7.5)***</td>
<td>1.60</td>
<td>54.07(9.2)***</td>
<td>59.71(8.2)***</td>
<td>65.12(7.5)***</td>
<td>1.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC$^a$</td>
<td>44.52</td>
<td>2.18,3253</td>
<td>&lt;.001</td>
<td>.03</td>
<td>55.86(10.1)</td>
<td>54.88(7.9)***</td>
<td>53.99(6.4)***</td>
<td>53.54(5.6)***</td>
<td>-.23</td>
<td>54.88(7.9)***</td>
<td>53.99(6.4)***</td>
<td>53.54(5.6)***</td>
<td>-.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC$^a$</td>
<td>1302.53</td>
<td>2.23,3324</td>
<td>&lt;.001</td>
<td>.47</td>
<td>52.81(10.2)</td>
<td>57.39(8.2)***</td>
<td>61.77(7.1)***</td>
<td>66.34(6.5)***</td>
<td>1.33</td>
<td>57.39(8.2)***</td>
<td>61.77(7.1)***</td>
<td>66.34(6.5)***</td>
<td>1.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Value</td>
<td>p-value</td>
<td>effect size</td>
<td>F-statistic</td>
<td>df1</td>
<td>df2</td>
<td>p-value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>---------</td>
<td>---------</td>
<td>-------------</td>
<td>-------------</td>
<td>-----</td>
<td>-----</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC6b</td>
<td>3964.64</td>
<td>&lt;.001</td>
<td>.73</td>
<td>55.40</td>
<td>1.15</td>
<td>2.10</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>67.98 (9.2)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>78.30 (9.4)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC7a</td>
<td>412.09</td>
<td>&lt;.001</td>
<td>.22</td>
<td>53.81 (11.8)</td>
<td>.22</td>
<td>.43</td>
<td>(6.2)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>56.42 (9.1)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>58.89 (7.0)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC8b</td>
<td>2159.89</td>
<td>&lt;.001</td>
<td>.59</td>
<td>54.39 (11.4)</td>
<td>.69</td>
<td>1.20</td>
<td>(7.1)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>62.22 (8.8)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>68.10 (7.5)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC9a</td>
<td>45.25</td>
<td>&lt;.001</td>
<td>.03</td>
<td>56.74 (11.4)</td>
<td>-.09</td>
<td>-.16</td>
<td>(5.8)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>55.67 (8.9)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>54.95 (7.0)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLSb</td>
<td>535.14</td>
<td>&lt;.001</td>
<td>.26</td>
<td>51.87 (10.2)</td>
<td>.34</td>
<td>.67</td>
<td>(8.3)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>55.32 (8.9)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>58.67 (8.3)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GICb</td>
<td>1264.98</td>
<td>&lt;.001</td>
<td>.46</td>
<td>52.92 (12.0)</td>
<td>.70</td>
<td>1.30</td>
<td>(9.9)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>61.26 (12.5)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>68.48 (11.4)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPCb</td>
<td>634.37</td>
<td>&lt;.001</td>
<td>.30</td>
<td>52.60 (11.1)</td>
<td>.37</td>
<td>.74</td>
<td>(8.0)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>56.72 (9.7)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>60.82 (8.6)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NUCb</td>
<td>1459.88</td>
<td>&lt;.001</td>
<td>.50</td>
<td>54.79 (11.6)</td>
<td>.64</td>
<td>1.14</td>
<td>(8.4)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>62.24 (9.6)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>68.01 (9.0)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COGb</td>
<td>582.07</td>
<td>&lt;.001</td>
<td>.28</td>
<td>56.35 (12.5)</td>
<td>.32</td>
<td>.63</td>
<td>(8.3)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>60.40 (10.3)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>64.23 (9.2)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUIb</td>
<td>1698.93</td>
<td>&lt;.001</td>
<td>.53</td>
<td>49.63 (11.5)</td>
<td>1.06</td>
<td>1.98</td>
<td>(13.7)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>61.87 (15.5)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>72.42 (14.9)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>81.01 (13.7)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLP&lt;sup&gt;b&lt;/sup&gt;</td>
<td>611.72</td>
<td>2.49,3706</td>
<td>&lt;.001</td>
<td>.29</td>
<td>50.28 (10.9)</td>
<td>50.28 (10.7)***</td>
<td>.00</td>
<td>59.43 (10.6)***</td>
<td>.84</td>
<td>63.52 (10.2)***</td>
<td>1.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFD&lt;sup&gt;b&lt;/sup&gt;</td>
<td>87.03</td>
<td>2.34,3479</td>
<td>&lt;.001</td>
<td>.06</td>
<td>53.20 (11.9)</td>
<td>54.72 (10.0)***</td>
<td>.13</td>
<td>55.98 (8.3)***</td>
<td>.23</td>
<td>57.53 (7.6)***</td>
<td>.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFC&lt;sup&gt;a&lt;/sup&gt;</td>
<td>54.42</td>
<td>2.18,3241</td>
<td>&lt;.001</td>
<td>.04</td>
<td>53.67 (11.0)</td>
<td>54.47 (8.7)***</td>
<td>.07</td>
<td>55.43 (7.2)***</td>
<td>.16</td>
<td>56.62 (6.9)***</td>
<td>.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STW&lt;sup&gt;b&lt;/sup&gt;</td>
<td>35.68</td>
<td>2.31,3437</td>
<td>&lt;.001</td>
<td>.02</td>
<td>52.71 (11.0)</td>
<td>53.27 (9.5)***</td>
<td>.05</td>
<td>54.33 (8.5)***</td>
<td>.15</td>
<td>55.28 (8.1)***</td>
<td>.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AXY&lt;sup&gt;b&lt;/sup&gt;</td>
<td>662.56</td>
<td>2.54,3780</td>
<td>&lt;.001</td>
<td>.31</td>
<td>56.08 (13.7)</td>
<td>62.28 (13.4)***</td>
<td>.45</td>
<td>67.87 (12.8)***</td>
<td>.86</td>
<td>73.14 (12.1)***</td>
<td>1.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANP&lt;sup&gt;b&lt;/sup&gt;</td>
<td>98.41</td>
<td>2.27</td>
<td>&lt;.001</td>
<td>.06</td>
<td>52.80 (11.2)</td>
<td>54.34 (8.9)***</td>
<td>.14</td>
<td>55.56 (7.4)***</td>
<td>.25</td>
<td>56.90 (6.9)***</td>
<td>.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRF&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1848.46</td>
<td>2.63,3913</td>
<td>&lt;.001</td>
<td>.55</td>
<td>53.38 (11.8)</td>
<td>62.81 (11.7)***</td>
<td>.80</td>
<td>71.41 (11.4)***</td>
<td>1.53</td>
<td>79.77 (11.1)***</td>
<td>2.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSF&lt;sup&gt;b&lt;/sup&gt;</td>
<td>216.89</td>
<td>2.29,3411</td>
<td>&lt;.001</td>
<td>.13</td>
<td>48.14 (8.6)</td>
<td>49.67 (7.0)***</td>
<td>.18</td>
<td>51.39 (8.6)***</td>
<td>.38</td>
<td>52.91 (5.9)***</td>
<td>.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JCP&lt;sup&gt;b&lt;/sup&gt;</td>
<td>647.17</td>
<td>2.53,3774</td>
<td>&lt;.001</td>
<td>.30</td>
<td>50.50 (10.3)</td>
<td>54.63 (9.4)***</td>
<td>.40</td>
<td>58.46 (8.5)***</td>
<td>.77</td>
<td>62.24 (8.2)***</td>
<td>1.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUB&lt;sup&gt;b&lt;/sup&gt;</td>
<td>491.32</td>
<td>2.34,3487</td>
<td>&lt;.001</td>
<td>.25</td>
<td>52.82 (11.9)</td>
<td>56.61 (10.4)***</td>
<td>.32</td>
<td>60.58 (10.0)***</td>
<td>.65</td>
<td>63.97 (9.6)***</td>
<td>.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>------</td>
<td>------</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGG</td>
<td>573.46</td>
<td>2.36,3542</td>
<td>&lt;.001</td>
<td>.28</td>
<td>51.90 (12.2)</td>
<td>55.86 (10.2)**</td>
<td>.32</td>
<td>59.67 (8.6)**</td>
<td>.64</td>
<td>63.31 (8.6)**</td>
<td>.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACT</td>
<td>.50</td>
<td>2.42,3609</td>
<td>.65</td>
<td>&lt;.01</td>
<td>51.64 (10.5)</td>
<td>55.05 (9.0)**</td>
<td>.32</td>
<td>58.40 (10.5)**</td>
<td>.64</td>
<td>61.93 (7.7)**</td>
<td>.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FML</td>
<td>560.95</td>
<td>2.34,3480</td>
<td>&lt;.001</td>
<td>.27</td>
<td>45.55 (7.9)</td>
<td>47.73 (6.8)**</td>
<td>.28</td>
<td>49.76 (6.1)**</td>
<td>.53</td>
<td>52.38 (6.4)**</td>
<td>.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPP</td>
<td>413.41</td>
<td>2.30,3424</td>
<td>&lt;.001</td>
<td>.22</td>
<td>45.54 (9.9)</td>
<td>49.25 (7.7)**</td>
<td>.37</td>
<td>52.06 (6.1)**</td>
<td>.66</td>
<td>55.28 (5.9)**</td>
<td>.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAV</td>
<td>751.72</td>
<td>2.18,3251</td>
<td>&lt;.001</td>
<td>.34</td>
<td>50.48 (10.8)</td>
<td>50.83 (8.3)*</td>
<td>.03</td>
<td>51.37 (6.5)**</td>
<td>.08</td>
<td>51.82 (5.7)**</td>
<td>.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHY</td>
<td>13.28</td>
<td>2.14,3192</td>
<td>&lt;.001</td>
<td>.01</td>
<td>50.73 (9.8)</td>
<td>59.88 (11.6)**</td>
<td>.93</td>
<td>68.07 (12.0)**</td>
<td>1.77</td>
<td>75.54 (12.2)**</td>
<td>2.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p ≤ .05, **p ≤ .01 level, ***p ≤ .001.

a – Greenhouse-Geisser, b – Huynh-Feldt
In the college sample, only the scores on one scale, ACT, were not significantly distorted across the levels of random response insertion. The results show that for ACT, unlike the other 36 scales, mean scale score level was not significantly affected by the level of random response insertion \( F(2.42, 3609) = .50, \ p > .05 \). In contrast, three scales (THD, RC6, and SUI) showed a large effect even at 30% random responses inserted, with scale scores rapidly inflating across increasing levels of random responding. Overall, 25 out of the 37 scales examined (all of the Higher Order Scales, 7 of the Restructured Clinical Scales, and 15 of the Specific Problems Scales) were distorted more than 5 T Score points in the 90% random responding condition.

Repeated Measures ANOVAs are reported for the outpatient sample in Table 11. Results for EID were not significant, so this scale was excluded from further analyses. In addition, Repeated Measures ANOVAs for MLS, ANP and SAV were similarly not significant. Results of the Repeated Measures ANOVA indicate that THD scores were distorted as a result of random response insertion \( F(2.51, 2257) = 2449.25, \ p < .001 \). THD scores became increasingly distorted under each successive random response insertion condition, and mean scores in the 30, 60 and 90% random response insertion conditions were statistically significantly different from the zero percent random response insertion condition. These differences reached a large effect size for all three comparisons \( (d = .96, 1.78, \text{ and } 2.54, \text{ respectively}) \). RC6 was the only other scale that showed such a dramatic effect on mean scale scores in the outpatient sample. Most other scales in this sample showed a more moderate pattern of increasing scale score distortion, such as the one exhibited by BXD. Although the mean scores at 30, 60 and 90% random
Table 11. The Effects of 30%, 60%, and 90% Random Response Insertion on MMPI-2-RF Scale Scores in an Outpatient Sample

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>$\eta^2_{\text{PARTIAL}}$</th>
<th>None Mean(SD)</th>
<th>30% Mean(SD)</th>
<th>60% Mean(SD)</th>
<th>90% Mean(SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>EID</td>
<td>4.01</td>
<td>1.48, 1333</td>
<td>.03</td>
<td>&lt;.01</td>
<td>56.43(12.1)</td>
<td>68.00(9.5)**</td>
<td>77.96(9.2)**</td>
<td>87.12(8.1)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.51, 2257</td>
<td>&lt;.001</td>
<td>.73</td>
<td>55.60(11.6)</td>
<td>58.10(9.1)**</td>
<td>60.41(7.5)**</td>
<td>63.20(6.3)**</td>
</tr>
<tr>
<td></td>
<td>222.17</td>
<td>2.03, 1821</td>
<td>&lt;.001</td>
<td>.20</td>
<td>65.72(13.3)</td>
<td>65.36(9.5)*</td>
<td>64.69(6.6)**</td>
<td>63.96(4.6)**</td>
</tr>
<tr>
<td></td>
<td>13.85</td>
<td>1.54, 1380</td>
<td>&lt;.001</td>
<td>.02</td>
<td>63.22(14.8)</td>
<td>67.51(10.4)**</td>
<td>71.14(7.2)**</td>
<td>74.15(6.2)**</td>
</tr>
<tr>
<td></td>
<td>367.11</td>
<td>1.74, 1567</td>
<td>&lt;.001</td>
<td>.29</td>
<td>62.51(11.3)</td>
<td>63.83(12.0)**</td>
<td>65.03(9.6)**</td>
<td>66.83(7.9)**</td>
</tr>
<tr>
<td></td>
<td>42.65</td>
<td>2.00, 1797</td>
<td>&lt;.001</td>
<td>.05</td>
<td>55.97(11.3)</td>
<td>54.92(8.6)**</td>
<td>54.24(7.1)**</td>
<td>53.61(5.8)**</td>
</tr>
<tr>
<td></td>
<td>23.17</td>
<td>2.04, 1833</td>
<td>&lt;.001</td>
<td>.03</td>
<td>59.50(11.7)</td>
<td>62.02(9.3)**</td>
<td>64.51(7.6)**</td>
<td>66.99(6.6)**</td>
</tr>
<tr>
<td></td>
<td>207.87</td>
<td>2.09, 1879</td>
<td>&lt;.001</td>
<td>.19</td>
<td>62.51(11.7)</td>
<td>62.02(9.3)**</td>
<td>64.51(7.6)**</td>
<td>66.99(6.6)**</td>
</tr>
<tr>
<td>Location</td>
<td>Value</td>
<td>p-value</td>
<td>Engine</td>
<td>Value</td>
<td>p-value</td>
<td>Engine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>---------</td>
<td>--------</td>
<td>-------</td>
<td>---------</td>
<td>--------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC6b</td>
<td>1605.85</td>
<td>&lt;.001</td>
<td>.64</td>
<td>60.19</td>
<td>13.5</td>
<td>(10.3)*** .83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC7a</td>
<td>53.29</td>
<td>&lt;.001</td>
<td>.06</td>
<td>58.40</td>
<td>12.6</td>
<td>(9.5)*** .10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC8b</td>
<td>1102.58</td>
<td>&lt;.001</td>
<td>.55</td>
<td>55.60</td>
<td>11.8</td>
<td>(9.1)*** .62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC9a</td>
<td>47.27</td>
<td>&lt;.001</td>
<td>.05</td>
<td>50.49</td>
<td>10.1</td>
<td>(8.1)*** .09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLSa</td>
<td>1.74</td>
<td>&lt;.17</td>
<td>&lt;.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GICb</td>
<td>215.27</td>
<td>&lt;.001</td>
<td>.19</td>
<td>62.83</td>
<td>17.3</td>
<td>(14.0)*** .32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPCb</td>
<td>51.61</td>
<td>&lt;.001</td>
<td>.05</td>
<td>60.39</td>
<td>13.8</td>
<td>(10.9)*** .12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NUCa</td>
<td>352.38</td>
<td>&lt;.001</td>
<td>.28</td>
<td>60.38</td>
<td>14.6</td>
<td>(11.2)*** .38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COGa</td>
<td>116.31</td>
<td>&lt;.001</td>
<td>.12</td>
<td>61.23</td>
<td>15.1</td>
<td>(11.5)*** .19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUib</td>
<td>517.46</td>
<td>&lt;.001</td>
<td>.37</td>
<td>58.12</td>
<td>18.6</td>
<td>(17.6)*** .45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GIC</td>
<td>215.27</td>
<td>&lt;.001</td>
<td>.19</td>
<td>62.83</td>
<td>17.3</td>
<td>(14.0)*** .32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPC</td>
<td>51.61</td>
<td>&lt;.001</td>
<td>.05</td>
<td>60.39</td>
<td>13.8</td>
<td>(10.9)*** .12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NUC</td>
<td>352.38</td>
<td>&lt;.001</td>
<td>.28</td>
<td>60.38</td>
<td>14.6</td>
<td>(11.2)*** .38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG</td>
<td>116.31</td>
<td>&lt;.001</td>
<td>.12</td>
<td>61.23</td>
<td>15.1</td>
<td>(11.5)*** .19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUI</td>
<td>517.46</td>
<td>&lt;.001</td>
<td>.37</td>
<td>58.12</td>
<td>18.6</td>
<td>(17.6)*** .45</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFD</td>
<td>80.59</td>
<td>2.37</td>
<td>2132</td>
<td>&lt;.001</td>
<td>.08</td>
<td>58.07</td>
<td>(14.7)</td>
<td>60.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AXY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JCP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Value</td>
<td>Mean Difference</td>
<td>p-value</td>
<td>Standard Error</td>
<td>Confidence Interval</td>
<td>Significance</td>
<td>Mean</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>-----</td>
<td>-------</td>
<td>-----------------</td>
<td>---------</td>
<td>----------------</td>
<td>---------------------</td>
<td>--------------</td>
<td>------</td>
<td>---------------------</td>
</tr>
<tr>
<td>AGGb</td>
<td>221.51</td>
<td>.20</td>
<td>&lt;.001</td>
<td>54.34</td>
<td>(12.0)</td>
<td>.27</td>
<td>60.18</td>
<td>(9.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>63.70</td>
<td>(8.5)</td>
</tr>
<tr>
<td>ACTb</td>
<td>23.22</td>
<td>.03</td>
<td>&lt;.001</td>
<td>51.02</td>
<td>(11.2)</td>
<td>.03</td>
<td>52.90</td>
<td>(9.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>53.71</td>
<td>(8.5)</td>
</tr>
<tr>
<td>FMLa</td>
<td>22.81</td>
<td>.03</td>
<td>&lt;.001</td>
<td>60.03</td>
<td>(13.6)</td>
<td>.04</td>
<td>62.00</td>
<td>(8.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>62.86</td>
<td>(8.2)</td>
</tr>
<tr>
<td>IPPa</td>
<td>6.51</td>
<td>.01</td>
<td>.001</td>
<td>51.69</td>
<td>(11.5)</td>
<td>.01</td>
<td>52.03</td>
<td>(7.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>52.95</td>
<td>(6.4)</td>
</tr>
<tr>
<td>SAVa</td>
<td>.51</td>
<td>1.96,1757</td>
<td>.59</td>
<td>&lt;.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHYa</td>
<td>8.66</td>
<td>.01</td>
<td>&lt;.001</td>
<td>53.44</td>
<td>(11.3)</td>
<td>-.06</td>
<td>52.35</td>
<td>(6.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>51.98</td>
<td>(5.9)</td>
</tr>
<tr>
<td>DSFB</td>
<td>565.91</td>
<td>.39</td>
<td>&lt;.001</td>
<td>53.75</td>
<td>(13.2)</td>
<td>.65</td>
<td>68.98</td>
<td>(13.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>74.82</td>
<td>(12.4)</td>
</tr>
</tbody>
</table>

** - p ≤ .05, *** - p ≤ .01 level, *** - p ≤ .001.

a Greenhouse Geisser; b Huynh-Feldt
responses insertion were statistically significantly different from the 0% random response insertion condition, the effect was small at the 30 \( (d = .22) \) and 60\% \( (d = .41) \) random response insertion conditions, and moderate in the 90\% \( (d = .61) \) random response insertion condition. Mean score differences did not reach a small effect size for four scales even when 90\% of responses were replaced with random ones in this sample. Most interestingly, RC3, SFD, and STW showed a negative effect in this sample. For these three scales, mean scores actually significantly decreased as greater amounts of random responses were inserted. In each case, the mean score on the scale in the 0\% random response insertion condition was above 50T (i.e., mean scores were 55.97 \[11.3\] for RC3; 62.45 \[12.8\] for SFD, and 60.15 \[12.2\] for STW). In this sample, 17 of the 37 scales (2 of the Higher Order Scales, 5 of the Restructured Clinical Scales, and 10 of the Specific Problems Scales) were distorted more than 5 T Score points in the 90\% random responding condition.

Table 12 reports the results for the inpatient sample. The Repeated Measures ANOVA for EID was not significant in this sample either. In addition, Repeated Measures ANOVAs for RC3, NFC, and SAV were not significant; so these scales were excluded from further analyses. However, the Repeated Measures ANOVA for THD was significant, indicating THD scores were distorted as a result of random response insertion \( (F(2.33,4449) = 4693.42, \ p < .001) \). As can be seen in the table, THD the mean scores in the 30, 60 and 90 \% random response insertion conditions were statistically significantly different from the zero percent random response condition, and those differences reached a large effect size for all three comparisons \( (d = .84, \ d = 1.57, \text{ and } d = 2.26, \text{ respectively}) \).
Table 12. The Effects of 30%, 60%, and 90% Random Response Insertion on MMPI-2-RF Scale Scores in an Inpatient Sample

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>( \eta^2_{\text{PARTIAL}} )</th>
<th>None Mean(SD)</th>
<th>30% Mean(SD)</th>
<th>60% Mean(SD)</th>
<th>90% Mean(SD)</th>
<th>d</th>
<th>60% d</th>
<th>90% d</th>
</tr>
</thead>
<tbody>
<tr>
<td>EID</td>
<td>.60</td>
<td>1.44,2752</td>
<td>.50</td>
<td>&lt;.01</td>
<td>56.43 (13.6)</td>
<td>67.90 (10.5)***</td>
<td>77.82 (9.3)***</td>
<td>87.23 (8.2)***</td>
<td>.84</td>
<td>1.5</td>
<td>2.26</td>
</tr>
<tr>
<td>THD</td>
<td>4693.42</td>
<td>2.33,4449</td>
<td>&lt;.001</td>
<td>.71</td>
<td>55.93 (12.1)</td>
<td>58.29 (9.5)***</td>
<td>60.80 (7.4)***</td>
<td>62.93 (6.2)***</td>
<td>.20</td>
<td>.40</td>
<td>.58</td>
</tr>
<tr>
<td>BXD</td>
<td>395.63</td>
<td>1.98,3794</td>
<td>&lt;.001</td>
<td>.17</td>
<td>63.30 (16.0)</td>
<td>64.19 (12.6)***</td>
<td>65.58 (9.6)***</td>
<td>66.27 (7.8)***</td>
<td>.06</td>
<td>.14</td>
<td>.19</td>
</tr>
<tr>
<td>RCd</td>
<td>5.46</td>
<td>1.50,2861</td>
<td>.01</td>
<td>&lt;.01</td>
<td>5.46 (1.50)</td>
<td>5.46 (1.50)</td>
<td>5.46 (1.50)</td>
<td>5.46 (1.50)</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>RC1</td>
<td>1292.24</td>
<td>1.89,3616</td>
<td>&lt;.001</td>
<td>.40</td>
<td>60.43 (13.9)</td>
<td>65.77 (9.8)***</td>
<td>70.09 (7.3)***</td>
<td>74.03 (6.0)***</td>
<td>.38</td>
<td>.69</td>
<td>.98</td>
</tr>
<tr>
<td>RC2</td>
<td>46.34</td>
<td>1.95,3727</td>
<td>&lt;.001</td>
<td>.02</td>
<td>39.30 (16.0)</td>
<td>44.19 (12.6)***</td>
<td>45.87 (9.6)***</td>
<td>48.47 (7.8)***</td>
<td>.06</td>
<td>.14</td>
<td>.19</td>
</tr>
<tr>
<td>RC3</td>
<td>6.69</td>
<td>1.97,3764</td>
<td>.001</td>
<td>&lt;.01</td>
<td>6.69 (1.97)</td>
<td>6.69 (1.97)</td>
<td>6.69 (1.97)</td>
<td>6.69 (1.97)</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>Value 1</td>
<td>Value 2</td>
<td>p-value</td>
<td>Value 3</td>
<td>Value 4</td>
<td>Value 5</td>
<td>Value 6</td>
<td>Value 7</td>
<td>Value 8</td>
<td>Value 9</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>---------</td>
<td>---------</td>
<td>----------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>RC4</td>
<td>381.53</td>
<td>1.99,3791</td>
<td>&lt;.001</td>
<td>59.57</td>
<td>(12.9)</td>
<td>62.14</td>
<td>(10.2)***</td>
<td>.20</td>
<td>64.44</td>
<td>(7.9)***</td>
<td>.38</td>
</tr>
<tr>
<td>RC6</td>
<td>3247.80</td>
<td>2.42,4629</td>
<td>&lt;.001</td>
<td>60.41</td>
<td>(14.5)</td>
<td>71.46</td>
<td>(11.1)***</td>
<td>.76</td>
<td>79.89</td>
<td>(9.8)***</td>
<td>1.3</td>
</tr>
<tr>
<td>RC7</td>
<td>204.51</td>
<td>1.80,3439</td>
<td>&lt;.001</td>
<td>56.59</td>
<td>(13.3)</td>
<td>58.30</td>
<td>(10.0)***</td>
<td>.13</td>
<td>60.07</td>
<td>(7.5)***</td>
<td>.26</td>
</tr>
<tr>
<td>RC8</td>
<td>2074.98</td>
<td>2.22,4243</td>
<td>&lt;.001</td>
<td>56.21</td>
<td>(12.6)</td>
<td>63.14</td>
<td>(9.6)***</td>
<td>.55</td>
<td>68.89</td>
<td>(8.0)***</td>
<td>1.0</td>
</tr>
<tr>
<td>RC9</td>
<td>118.30</td>
<td>1.98,3777</td>
<td>&lt;.001</td>
<td>49.93</td>
<td>(11.3)</td>
<td>50.98</td>
<td>(8.6)***</td>
<td>.09</td>
<td>52.36</td>
<td>(6.7)***</td>
<td>.22</td>
</tr>
<tr>
<td>MLS</td>
<td>17.23</td>
<td>2.11,4039</td>
<td>&lt;.001</td>
<td>65.23</td>
<td>(14.2)</td>
<td>64.64</td>
<td>(11.6)***</td>
<td>-.04</td>
<td>63.70</td>
<td>(9.3)***</td>
<td>-.11</td>
</tr>
<tr>
<td>GIC</td>
<td>711.17</td>
<td>2.54,4864</td>
<td>&lt;.001</td>
<td>59.52</td>
<td>(15.5)</td>
<td>65.71</td>
<td>(13.3)***</td>
<td>.40</td>
<td>70.51</td>
<td>(11.3)***</td>
<td>.71</td>
</tr>
<tr>
<td>HPC</td>
<td>315.87</td>
<td>2.30,4395</td>
<td>&lt;.001</td>
<td>56.26</td>
<td>(12.7)</td>
<td>59.60</td>
<td>(10.6)***</td>
<td>.26</td>
<td>61.88</td>
<td>(8.8)***</td>
<td>.44</td>
</tr>
<tr>
<td>NUC</td>
<td>769.08</td>
<td>2.35,4491</td>
<td>&lt;.001</td>
<td>60.93</td>
<td>(14.3)</td>
<td>65.98</td>
<td>(11.4)***</td>
<td>.35</td>
<td>70.06</td>
<td>(9.4)***</td>
<td>.64</td>
</tr>
<tr>
<td>COG</td>
<td>236.60</td>
<td>2.12,4045</td>
<td>&lt;.001</td>
<td>61.02</td>
<td>(15.1)</td>
<td>63.73</td>
<td>(11.8)***</td>
<td>.18</td>
<td>65.89</td>
<td>(9.5)***</td>
<td>.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUI&lt;sup&gt;b&lt;/sup&gt;</td>
<td>367.32</td>
<td>2.43,4636</td>
<td>&lt;.001</td>
<td>.16</td>
<td>68.48</td>
<td>74.28</td>
<td>78.94</td>
<td>82.42</td>
<td>.27</td>
<td>(21.3)</td>
<td>(17.9)***</td>
</tr>
<tr>
<td>HLP&lt;sup&gt;b&lt;/sup&gt;</td>
<td>121.51</td>
<td>2.32,4435</td>
<td>&lt;.001</td>
<td>.06</td>
<td>58.59</td>
<td>60.68</td>
<td>62.28</td>
<td>64.41</td>
<td>.14</td>
<td>(15.2)</td>
<td>(12.8)***</td>
</tr>
<tr>
<td>SFD&lt;sup&gt;b&lt;/sup&gt;</td>
<td>39.03</td>
<td>2.28,4364</td>
<td>&lt;.001</td>
<td>.02</td>
<td>60.53</td>
<td>59.74</td>
<td>58.51</td>
<td>57.96</td>
<td>-.06</td>
<td>(12.9)</td>
<td>(10.7)***</td>
</tr>
<tr>
<td>NFC&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.51</td>
<td>2.04,3894</td>
<td>.61</td>
<td>&lt;.01</td>
<td>57.83</td>
<td>56.78</td>
<td>56.21</td>
<td>55.89</td>
<td>-.09</td>
<td>(12.1)</td>
<td>(10.4)***</td>
</tr>
<tr>
<td>STW&lt;sup&gt;b&lt;/sup&gt;</td>
<td>21.97</td>
<td>2.26,4318</td>
<td>&lt;.001</td>
<td>.01</td>
<td>62.21</td>
<td>66.43</td>
<td>70.64</td>
<td>73.51</td>
<td>.25</td>
<td>(16.8)</td>
<td>(14.6)***</td>
</tr>
<tr>
<td>AXY&lt;sup&gt;b&lt;/sup&gt;</td>
<td>346.53</td>
<td>2.43,4654</td>
<td>&lt;.001</td>
<td>.15</td>
<td>54.32</td>
<td>55.53</td>
<td>56.12</td>
<td>57.10</td>
<td>-.10</td>
<td>(11.8)</td>
<td>(9.5)***</td>
</tr>
<tr>
<td>ANP&lt;sup&gt;a&lt;/sup&gt;</td>
<td>51.92</td>
<td>2.14,4094</td>
<td>&lt;.001</td>
<td>.03</td>
<td>54.72</td>
<td>56.74</td>
<td>71.71</td>
<td>79.77</td>
<td>.73</td>
<td>(12.3)</td>
<td>(11.9)***</td>
</tr>
<tr>
<td>BRF&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2133.84</td>
<td>2.59,4956</td>
<td>&lt;.001</td>
<td>.53</td>
<td>48.98</td>
<td>50.22</td>
<td>51.45</td>
<td>52.98</td>
<td>.13</td>
<td>(9.3)</td>
<td>(7.4)***</td>
</tr>
<tr>
<td>MSF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>174.35</td>
<td>2.25,4305</td>
<td>&lt;.001</td>
<td>.08</td>
<td>57.87</td>
<td>59.50</td>
<td>61.34</td>
<td>62.73</td>
<td>.13</td>
<td>(13.0)</td>
<td>(11.0)***</td>
</tr>
<tr>
<td>Variable</td>
<td>Value Mean</td>
<td>Value SD</td>
<td>Score Mean</td>
<td>Score SD</td>
<td>Score</td>
<td>Score SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
<td>----------</td>
<td>------------</td>
<td>----------</td>
<td>-------</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUB</td>
<td>302.05</td>
<td>2.20</td>
<td>56.39 (14.0)</td>
<td>.14</td>
<td>61.77 (10.1)</td>
<td>.38 (9.7)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGG</td>
<td>592.69</td>
<td>2.32</td>
<td>52.61 (12.3)</td>
<td>.24</td>
<td>59.69 (8.9)***</td>
<td>.58 (8.5)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACT</td>
<td>54.94</td>
<td>2.23</td>
<td>50.40 (12.5)</td>
<td>.03</td>
<td>52.40 (8.8)***</td>
<td>.16 (8.4)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FML</td>
<td>240.92</td>
<td>2.18</td>
<td>55.84 (13.2)</td>
<td>.11</td>
<td>60.07 (8.8)***</td>
<td>.32 (8.3)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPP</td>
<td>24.72</td>
<td>2.15</td>
<td>51.11 (11.4)</td>
<td>.01</td>
<td>52.00 (7.1)***</td>
<td>.08 (6.3)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAV</td>
<td>.56</td>
<td>1.96</td>
<td>56.39 (14.4)</td>
<td>.38</td>
<td>63.32 (13.6)***</td>
<td>.52 (12.9)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHY</td>
<td>20.87</td>
<td>2.16</td>
<td>53.33 (11.3)</td>
<td>.01</td>
<td>52.10 (6.4)***</td>
<td>-.11 (5.5)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSF</td>
<td>1146.45</td>
<td>2.57</td>
<td>55.83 (14.4)</td>
<td>.38</td>
<td>70.27 (12.9)***</td>
<td>1.0 (11.9)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* - p \leq .05, ** - p \leq .01 level, *** - p \leq .001.

a – Greenhouse-Geisser; b – Huynh-Feldt
No other scale was affected so dramatically in this sample. Most other scales in the inpatient sample demonstrated a more moderate, but significant pattern of increasing score distortion, although differences in mean scores did not reach a small effect size for five scales even when 90% of responses were replaced with random ones. One scale (SFD) showed a small negative effect in this sample. Nineteen of the 37 scales (2 Higher Order Scales, 5 Restructured Clinical Scales, and 12 Specific Problems Scales) were distorted by 5 T-Score points or greater in the 90% random response insertion condition.

Results for the criminal forensic sample are reported in Table 13. A similar pattern of results was found for this sample, as THD exhibited the most rapid and extreme scale score distortion. Nevertheless, the majority of MMPI-2-RF Scales demonstrated moderate scale score distortion under increasing proportions of random response insertion. In this sample, only the Repeated Measures ANOVA for RC3 was non-significant, and only two scales (STW and SHY) did not exhibit even a small effect size at 90% random response insertion. In addition, 23 scales (all of the Higher Order Scales, 8 of the Restructured Clinical Scales, and 12 of the Specific Problems Scales) displayed mean scores that were distorted 5 T-Score points or more in the 90% random responding condition.

Table 14 reports the results of the Repeated Measures ANOVAs, post-hoc tests, and associated effect sizes for the civil forensic sample. The Repeated Measures ANOVAs were not significant for three scales, (SFD, STW, and SHY) in this sample. Random response insertion had a large effect on both THD and RC6 Scale scores ($d = 1.08$; $d = 1.07$), even at 30% random response insertion. Taken as a whole, the other
Table 13. The Effects of 30%, 60%, and 90% Random Response Insertion on MMPI-2-RF Scale Scores in a Criminal Forensic Sample

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>(\eta^2_{\text{PARTIAL}})</th>
<th>None</th>
<th>30%</th>
<th>60%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean(SD)</td>
<td>Mean(SD)</td>
<td>d</td>
<td>Mean(SD)</td>
<td>d</td>
<td>Mean(SD)</td>
<td>d</td>
<td>Mean(SD)</td>
</tr>
<tr>
<td>EID(^a)</td>
<td>247.41</td>
<td>1.52,1513</td>
<td>&lt;.001</td>
<td>.20</td>
<td>56.26 (13.8)</td>
<td>59.09 (9.4)**</td>
<td>.21</td>
<td>61.16 (6.5)**</td>
</tr>
<tr>
<td>THD(^b)</td>
<td>2635.02</td>
<td>2.39,2379</td>
<td>&lt;.001</td>
<td>.73</td>
<td>55.86 (13.6)</td>
<td>67.42 (10.4)**</td>
<td>.85</td>
<td>77.48 (9.3)**</td>
</tr>
<tr>
<td>BXD(^a)</td>
<td>149.28</td>
<td>1.91,1895</td>
<td>&lt;.001</td>
<td>.13</td>
<td>57.43 (12.2)</td>
<td>59.55 (9.2)**</td>
<td>.17</td>
<td>61.15 (7.3)**</td>
</tr>
<tr>
<td>RC(^d)</td>
<td>175.62</td>
<td>1.60,1585</td>
<td>&lt;.001</td>
<td>.15</td>
<td>56.98 (13.2)</td>
<td>59.71 (8.9)**</td>
<td>.21</td>
<td>61.33 (6.3)**</td>
</tr>
<tr>
<td>RC(^1)</td>
<td>1030.14</td>
<td>1.91,1901</td>
<td>&lt;.001</td>
<td>.51</td>
<td>56.99 (13.8)</td>
<td>63.50 (9.5)**</td>
<td>.47</td>
<td>68.75 (7.1)**</td>
</tr>
<tr>
<td>RC(^2)</td>
<td>398.32</td>
<td>2.19,2173</td>
<td>&lt;.001</td>
<td>.29</td>
<td>54.47 (13.4)</td>
<td>58.13 (10.8)**</td>
<td>.27</td>
<td>61.79 (9.1)**</td>
</tr>
<tr>
<td>RC(^3)</td>
<td>.62</td>
<td>1.91,1901</td>
<td>.53</td>
<td>&lt;.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC(^4)</td>
<td>174.91</td>
<td>2.00,1990</td>
<td>&lt;.001</td>
<td>.15</td>
<td>60.48 (12.7)</td>
<td>62.84 (9.8)**</td>
<td>.19</td>
<td>64.33 (7.9)**</td>
</tr>
<tr>
<td></td>
<td>Value</td>
<td>p-value</td>
<td>Mean</td>
<td>Standard Deviation</td>
<td>t-statistic</td>
<td>Degrees of Freedom</td>
<td>p-value</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>---------</td>
<td>------</td>
<td>--------------------</td>
<td>-------------</td>
<td>--------------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>RC6b</td>
<td>1681.22</td>
<td>&lt;.001</td>
<td>60.03</td>
<td>14.7</td>
<td>.73</td>
<td>79.29 (9.9)</td>
<td>.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>70.76</td>
<td>11.3</td>
<td>.73</td>
<td>87.84 (8.7)</td>
<td>1.89</td>
<td></td>
</tr>
<tr>
<td>RC7a</td>
<td>332.32</td>
<td>&lt;.001</td>
<td>52.64</td>
<td>12.3</td>
<td>.26</td>
<td>58.53 (6.9)</td>
<td>.48</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>55.78</td>
<td>9.11</td>
<td>.26</td>
<td>61.14 (5.9)</td>
<td>.69</td>
<td></td>
</tr>
<tr>
<td>RC8a</td>
<td>1180.77</td>
<td>&lt;.001</td>
<td>55.21</td>
<td>12.9</td>
<td>.57</td>
<td>68.63 (8.2)</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>62.62</td>
<td>9.7</td>
<td>.57</td>
<td>74.39 (7.4)</td>
<td>1.49</td>
<td></td>
</tr>
<tr>
<td>RC9a</td>
<td>69.20</td>
<td>&lt;.001</td>
<td>49.83</td>
<td>11.1</td>
<td>.12</td>
<td>52.16 (6.6)</td>
<td>.21</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>51.15</td>
<td>8.5</td>
<td>.12</td>
<td>53.56 (5.6)</td>
<td>.34</td>
<td></td>
</tr>
<tr>
<td>MLSa</td>
<td>62.68</td>
<td>&lt;.001</td>
<td>57.34</td>
<td>13.4</td>
<td>.15</td>
<td>61.00 (9.5)</td>
<td>.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>59.30</td>
<td>11.0</td>
<td>.15</td>
<td>62.11 (8.3)</td>
<td>.36</td>
<td></td>
</tr>
<tr>
<td>GICb</td>
<td>560.37</td>
<td>&lt;.001</td>
<td>55.90</td>
<td>14.7</td>
<td>.36</td>
<td>63.18 (13.4)</td>
<td>.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>63.18</td>
<td>13.4</td>
<td>.36</td>
<td>69.17 (14.7)</td>
<td>.90</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>69.17</td>
<td>14.7</td>
<td>.36</td>
<td>74.25 (10.0)</td>
<td>1.25</td>
<td></td>
</tr>
<tr>
<td>HPCb</td>
<td>256.96</td>
<td>&lt;.001</td>
<td>55.29</td>
<td>12.6</td>
<td>.21</td>
<td>58.67 (10.4)</td>
<td>.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>58.67</td>
<td>10.4</td>
<td>.21</td>
<td>61.90 (8.8)</td>
<td>.52</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>61.90</td>
<td>8.8</td>
<td>.21</td>
<td>64.88 (8.1)</td>
<td>.76</td>
<td></td>
</tr>
<tr>
<td>NUCb</td>
<td>643.60</td>
<td>&lt;.001</td>
<td>57.09</td>
<td>13.9</td>
<td>.39</td>
<td>63.24 (11.3)</td>
<td>.44</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>63.24</td>
<td>11.3</td>
<td>.39</td>
<td>68.48 (8.9)</td>
<td>.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>68.48</td>
<td>8.9</td>
<td>.39</td>
<td>73.09 (8.3)</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>COGa</td>
<td>414.74</td>
<td>&lt;.001</td>
<td>55.50</td>
<td>14.3</td>
<td>.29</td>
<td>60.43 (11.1)</td>
<td>.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>60.43</td>
<td>11.1</td>
<td>.29</td>
<td>64.26 (9.4)</td>
<td>.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>64.26</td>
<td>9.4</td>
<td>.29</td>
<td>68.11 (8.4)</td>
<td>.88</td>
<td></td>
</tr>
<tr>
<td>SUIb</td>
<td>636.24</td>
<td>&lt;.001</td>
<td>55.46</td>
<td>17.4</td>
<td>.39</td>
<td>65.77 (17.2)</td>
<td>.59</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>65.77</td>
<td>17.2</td>
<td>.39</td>
<td>73.62 (15.1)</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>73.62</td>
<td>15.1</td>
<td>.39</td>
<td>81.32 (13.2)</td>
<td>1.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>---</td>
<td>-----</td>
<td>---------</td>
<td>---</td>
<td>-----</td>
<td>---------</td>
<td>---</td>
<td>-----</td>
</tr>
<tr>
<td>HLP</td>
<td>284.36</td>
<td>2.47,2451</td>
<td>&lt;.001</td>
<td>.22</td>
<td>51.95 (12.9)</td>
<td>56.32 (11.5)***</td>
<td>.34</td>
<td>60.01 (10.7)***</td>
</tr>
<tr>
<td>SFD</td>
<td>21.24</td>
<td>2.32,2307</td>
<td>&lt;.001</td>
<td>.02</td>
<td>54.71 (12.7)</td>
<td>55.47 (10.2)***</td>
<td>.06</td>
<td>56.26 (8.5)***</td>
</tr>
<tr>
<td>NFC</td>
<td>53.18</td>
<td>2.10,2086</td>
<td>&lt;.001</td>
<td>.05</td>
<td>53.07 (11.6)</td>
<td>54.25 (9.4)***</td>
<td>.10</td>
<td>55.68 (7.5)***</td>
</tr>
<tr>
<td>STW</td>
<td>4.93</td>
<td>2.26,2244</td>
<td>.005</td>
<td>.01</td>
<td>54.11 (11.5)</td>
<td>54.39 (10.1)</td>
<td>.02</td>
<td>55.03 (8.7)**</td>
</tr>
<tr>
<td>AXY</td>
<td>361.70</td>
<td>2.51,2487</td>
<td>&lt;.001</td>
<td>.27</td>
<td>56.74 (15.3)</td>
<td>62.75 (13.8)***</td>
<td>.39</td>
<td>68.43 (13.0)***</td>
</tr>
<tr>
<td>ANP</td>
<td>95.15</td>
<td>2.13,2119</td>
<td>&lt;.001</td>
<td>.09</td>
<td>52.02 (12.1)</td>
<td>54.24 (9.4)***</td>
<td>.18</td>
<td>55.50 (7.4)**</td>
</tr>
<tr>
<td>BRF</td>
<td>1406.10</td>
<td>2.60,2584</td>
<td>&lt;.001</td>
<td>.59</td>
<td>52.21 (11.1)***</td>
<td>62.34 (11.4)***</td>
<td>.91</td>
<td>71.7 (11.3)***</td>
</tr>
<tr>
<td>MSF</td>
<td>73.99</td>
<td>2.18,2170</td>
<td>&lt;.001</td>
<td>.07</td>
<td>49.54 (9.4)</td>
<td>50.79 (7.3)***</td>
<td>.13</td>
<td>51.88 (6.1)***</td>
</tr>
<tr>
<td>JCP</td>
<td>46.24</td>
<td>2.28,2263</td>
<td>&lt;.001</td>
<td>.04</td>
<td>58.70 (13.3)</td>
<td>60.35 (11.1)***</td>
<td>.12</td>
<td>61.13 (9.2)***</td>
</tr>
<tr>
<td>SUB</td>
<td>132.92</td>
<td>2.37,2352</td>
<td>&lt;.001</td>
<td>.12</td>
<td>57.25 (14.2)</td>
<td>60.25 (12.0)***</td>
<td>.21</td>
<td>62.17 (10.0)***</td>
</tr>
<tr>
<td>Variable</td>
<td>Value</td>
<td>df</td>
<td>p</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>----</td>
<td>---</td>
<td>------</td>
<td>----</td>
<td>------</td>
<td>----</td>
<td>------</td>
</tr>
<tr>
<td>AGG^b</td>
<td>403.81</td>
<td>2.28,2265</td>
<td>&lt;.001</td>
<td>.29</td>
<td>51.47 (12.1)</td>
<td>.35</td>
<td>59.18 (8.4)***</td>
<td>.64</td>
</tr>
<tr>
<td>ACT^a</td>
<td>24.86</td>
<td>2.21,2195</td>
<td>&lt;.001</td>
<td>.02</td>
<td>50.83 (12.2)</td>
<td>.07</td>
<td>52.53 (9.1)***</td>
<td>.14</td>
</tr>
<tr>
<td>FML^a</td>
<td>292.87</td>
<td>2.14,2125</td>
<td>&lt;.001</td>
<td>.23</td>
<td>52.23 (13.1)</td>
<td>.24</td>
<td>58.52 (8.8)***</td>
<td>.48</td>
</tr>
<tr>
<td>IPP^a</td>
<td>63.53</td>
<td>2.20,2182</td>
<td>&lt;.001</td>
<td>.06</td>
<td>48.76 (10.2)</td>
<td>.14</td>
<td>51.29 (7.1)***</td>
<td>.25</td>
</tr>
<tr>
<td>SAV^a</td>
<td>39.13</td>
<td>2.04,2024</td>
<td>&lt;.001</td>
<td>.04</td>
<td>53.13 (11.8)</td>
<td>.08</td>
<td>54.89 (7.0)***</td>
<td>.15</td>
</tr>
<tr>
<td>SHY^a</td>
<td>9.38</td>
<td>2.14,2125</td>
<td>&lt;.001</td>
<td>.01</td>
<td>50.34 (10.8)</td>
<td>.03</td>
<td>50.69 (8.1)</td>
<td>.03</td>
</tr>
<tr>
<td>DSF^b</td>
<td>2270.95</td>
<td>2.64,2624</td>
<td>&lt;.001</td>
<td>.47</td>
<td>52.03 (12.3)</td>
<td>.68</td>
<td>68.26 (12.3)***</td>
<td>1.32</td>
</tr>
</tbody>
</table>

- *p ≤ .05, **p ≤ .01 level, ***p ≤ .001.
- ^a – Greenhouse-Geisser; ^b – Huynh-Feldt
Table 14. The Effects of 30%, 60%, and 90% Random Response Insertion on MMPI-2-RF Scale Scores in a Civil Forensic Sample

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>(\eta^2_{\text{PARTIAL}})</th>
<th>None Mean(SD)</th>
<th>30% Mean(SD)</th>
<th>d</th>
<th>60% Mean(SD)</th>
<th>d</th>
<th>90% Mean(SD)</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>EID</td>
<td>52.82</td>
<td>1.55,2354</td>
<td>&lt;.001</td>
<td>.03</td>
<td>61.35 (13.4)</td>
<td>62.51 (9.4)***</td>
<td>.09</td>
<td>63.24 (6.5)***</td>
<td>.14</td>
<td>64.03 (4.6)***</td>
<td>.20</td>
</tr>
<tr>
<td>THD</td>
<td>4813.38</td>
<td>2.52,3835</td>
<td>&lt;.001</td>
<td>.76</td>
<td>53.98 (11.2)</td>
<td>66.11 (9.1)***</td>
<td>1.08</td>
<td>76.71 (8.9)***</td>
<td>2.03</td>
<td>87.02 (8.2)***</td>
<td>2.95</td>
</tr>
<tr>
<td>BXD</td>
<td>1120.73</td>
<td>2.09,3180</td>
<td>&lt;.001</td>
<td>.42</td>
<td>49.62 (11.8)</td>
<td>54.15 (9.0)***</td>
<td>.38</td>
<td>58.36 (7.4)***</td>
<td>.74</td>
<td>62.41 (6.2)***</td>
<td>1.08</td>
</tr>
<tr>
<td>RCd</td>
<td>23.61</td>
<td>1.60,2424</td>
<td>&lt;.001</td>
<td>.02</td>
<td>61.87 (12.6)</td>
<td>62.73 (8.7)***</td>
<td>.07</td>
<td>63.24 (6.1)***</td>
<td>.11</td>
<td>63.57 (4.5)***</td>
<td>.13</td>
</tr>
<tr>
<td>RC1</td>
<td>94.64</td>
<td>1.81,2747</td>
<td>&lt;.001</td>
<td>.06</td>
<td>71.10 (14.1)</td>
<td>72.64 (10.3)***</td>
<td>.11</td>
<td>73.90 (7.6)***</td>
<td>.20</td>
<td>75.20 (6.1)***</td>
<td>.29</td>
</tr>
<tr>
<td>RC2</td>
<td>94.36</td>
<td>2.17,3294</td>
<td>&lt;.001</td>
<td>.06</td>
<td>61.96 (13.7)</td>
<td>63.45 (10.9)***</td>
<td>.11</td>
<td>65.16 (8.9)***</td>
<td>.23</td>
<td>66.47 (8.0)***</td>
<td>.33</td>
</tr>
<tr>
<td>RC3</td>
<td>88.48</td>
<td>2.02,3062</td>
<td>&lt;.001</td>
<td>.06</td>
<td>49.52</td>
<td>50.73</td>
<td>.11</td>
<td>51.52</td>
<td>.18</td>
<td>52.94</td>
<td>.31</td>
</tr>
<tr>
<td>Group</td>
<td>Value 1</td>
<td>Value 2</td>
<td>p-value</td>
<td>Value 3</td>
<td>Value 4</td>
<td>Value 5</td>
<td>Value 6</td>
<td>Value 7</td>
<td>Value 8</td>
<td>Value 9</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td><strong>RC4</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1487.20</td>
<td>2.22,3374</td>
<td>&lt;.001</td>
<td>.50</td>
<td>(10.9)</td>
<td>(8.2)***</td>
<td>(6.3)***</td>
<td>(5.5)***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50.85</td>
<td>(11.5)</td>
<td>56.28</td>
<td>(8.7)***</td>
<td>.47</td>
<td>61.05</td>
<td>(7.6)***</td>
</tr>
<tr>
<td><strong>RC6</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3824.98</td>
<td>2.69,4081</td>
<td>&lt;.001</td>
<td>.72</td>
<td>55.54</td>
<td>(12.1)</td>
<td>68.47</td>
<td>(9.5)***</td>
<td>1.07</td>
<td>78.10</td>
<td>(9.4)***</td>
</tr>
<tr>
<td><strong>RC7</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>251.37</td>
<td>1.90,2884</td>
<td>&lt;.001</td>
<td>.14</td>
<td>55.43</td>
<td>(12.8)</td>
<td>57.71</td>
<td>(9.7)***</td>
<td>.18</td>
<td>59.43</td>
<td>(7.2)***</td>
</tr>
<tr>
<td><strong>RC8</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2036.66</td>
<td>2.33,3535</td>
<td>&lt;.001</td>
<td>.57</td>
<td>54.85</td>
<td>(11.6)</td>
<td>62.02</td>
<td>(9.1)***</td>
<td>.62</td>
<td>68.31</td>
<td>(7.7)***</td>
</tr>
<tr>
<td><strong>RC9</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>295.45</td>
<td>2.05,3120</td>
<td>&lt;.001</td>
<td>.16</td>
<td>47.20</td>
<td>(10.5)</td>
<td>49.11</td>
<td>(8.0)***</td>
<td>.18</td>
<td>51.47</td>
<td>(6.6)***</td>
</tr>
<tr>
<td><strong>MLS</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>350.99</td>
<td>2.26,3442</td>
<td>&lt;.001</td>
<td>.19</td>
<td>72.53</td>
<td>(12.0)</td>
<td>69.61</td>
<td>(10.3)***</td>
<td>-.24</td>
<td>66.66</td>
<td>(8.8)***</td>
</tr>
<tr>
<td><strong>GIC</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>263.17</td>
<td>2.41,3670</td>
<td>&lt;.001</td>
<td>.15</td>
<td>64.83</td>
<td>(16.6)</td>
<td>69.11</td>
<td>(13.8)***</td>
<td>.26</td>
<td>72.17</td>
<td>(11.8)***</td>
</tr>
<tr>
<td><strong>HPC</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>20.77</td>
<td>2.30,3496</td>
<td>&lt;.001</td>
<td>.01</td>
<td>67.75</td>
<td>(12.4)</td>
<td>67.19</td>
<td>(10.3)***</td>
<td>-.05</td>
<td>66.37</td>
<td>(8.8)***</td>
</tr>
<tr>
<td><strong>NUC</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>118.11</td>
<td>2.20,3345</td>
<td>&lt;.001</td>
<td>.07</td>
<td>68.78</td>
<td>(14.6)</td>
<td>70.97</td>
<td>(11.6)***</td>
<td>.15</td>
<td>72.57</td>
<td>(9.3)***</td>
</tr>
<tr>
<td><strong>COG</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>50.47</td>
<td>2.12,3215</td>
<td>&lt;.001</td>
<td>.03</td>
<td>65.02</td>
<td>(5.5)***</td>
<td>66.26</td>
<td>(8.2)***</td>
<td>.08</td>
<td>67.92</td>
<td>(7.6)***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SUI</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1164.66</td>
<td>2.68,4072</td>
<td>&lt;.001</td>
<td>.43</td>
<td>54.64</td>
<td>(16.2)</td>
<td>.65</td>
<td>74.52</td>
<td>(14.9)</td>
<td>1.23</td>
<td>81.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HLP</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>149.45</td>
<td>2.43,3696</td>
<td>&lt;.001</td>
<td>.09</td>
<td>57.11</td>
<td>(14.3)</td>
<td>.17</td>
<td>61.83</td>
<td>(10.6)</td>
<td>.33</td>
<td>64.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SFD</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.56</td>
<td>2.31,3516</td>
<td>.59</td>
<td>&lt;.01</td>
<td>54.74</td>
<td>(12.3)</td>
<td>.06</td>
<td>55.79</td>
<td>(7.8)</td>
<td>.09</td>
<td>56.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NFC</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>24.31</td>
<td>2.15,3266</td>
<td>&lt;.001</td>
<td>.02</td>
<td>55.13</td>
<td>(12.2)</td>
<td>.06</td>
<td>56.44</td>
<td>(7.8)</td>
<td>.11</td>
<td>57.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>STW</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.28</td>
<td>2.33,3542</td>
<td>.09</td>
<td>&lt;.01</td>
<td>54.08</td>
<td>(12.0)</td>
<td>.80</td>
<td>71.88</td>
<td>(11.8)</td>
<td>1.48</td>
<td>79.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AXY</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>277.52</td>
<td>2.37,3598</td>
<td>&lt;.001</td>
<td>.15</td>
<td>62.23</td>
<td>(17.2)</td>
<td>.23</td>
<td>70.45</td>
<td>(13.0)</td>
<td>.48</td>
<td>73.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ANP</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>23.38</td>
<td>2.23,3390</td>
<td>&lt;.001</td>
<td>.02</td>
<td>55.13</td>
<td>(12.2)</td>
<td>.06</td>
<td>56.44</td>
<td>(7.8)</td>
<td>.11</td>
<td>57.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>BRF</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1721.03</td>
<td>2.57,3908</td>
<td>&lt;.001</td>
<td>.53</td>
<td>54.08</td>
<td>(12.0)</td>
<td>.80</td>
<td>71.88</td>
<td>(11.8)</td>
<td>1.48</td>
<td>79.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MSF</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>98.56</td>
<td>2.30,3499</td>
<td>&lt;.001</td>
<td>.06</td>
<td>49.63</td>
<td>(9.4)</td>
<td>.13</td>
<td>51.83</td>
<td>(6.2)</td>
<td>.23</td>
<td>52.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>JCP</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>376.12</td>
<td>2.40,3643</td>
<td>&lt;.001</td>
<td>.20</td>
<td>52.94</td>
<td>56.34</td>
<td>.27</td>
<td>59.32</td>
<td>.51</td>
<td>62.34</td>
<td>.76</td>
</tr>
</tbody>
</table>

**Note:**<sup>a</sup> Data obtained from blood samples,<sup>b</sup> data obtained from other sources.
<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SUB</td>
<td>1231</td>
<td>2.46,3744</td>
<td>&lt;.001</td>
<td>.45</td>
<td>47.78</td>
<td>(9.1)***</td>
<td>53.09</td>
<td>(8.8)***</td>
<td>.58</td>
</tr>
<tr>
<td>AGG</td>
<td>761.62</td>
<td>2.40,3650</td>
<td>&lt;.001</td>
<td>.33</td>
<td>50.12</td>
<td>(11.1)</td>
<td>54.86</td>
<td>(9.7)***</td>
<td>.43</td>
</tr>
<tr>
<td>ACT</td>
<td>112.13</td>
<td>2.40,3645</td>
<td>&lt;.001</td>
<td>.07</td>
<td>48.67</td>
<td>(10.8)</td>
<td>50.26</td>
<td>(9.4)***</td>
<td>.15</td>
</tr>
<tr>
<td>FML</td>
<td>630.56</td>
<td>2.34,3548</td>
<td>&lt;.001</td>
<td>.29</td>
<td>51.06</td>
<td>(11.2)</td>
<td>54.54</td>
<td>(9.5)***</td>
<td>.31</td>
</tr>
<tr>
<td>IPP</td>
<td>112.24</td>
<td>2.31,3514</td>
<td>&lt;.001</td>
<td>.07</td>
<td>48.75</td>
<td>(9.7)</td>
<td>49.96</td>
<td>(8.1)***</td>
<td>.12</td>
</tr>
<tr>
<td>SAV</td>
<td>16.67</td>
<td>2.03,3086</td>
<td>&lt;.001</td>
<td>.01</td>
<td>54.23</td>
<td>(12.2)</td>
<td>55.13</td>
<td>(9.2)***</td>
<td>.07</td>
</tr>
<tr>
<td>SHY</td>
<td>2.60</td>
<td>2.21,3354</td>
<td>.07</td>
<td>&lt;.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSF</td>
<td>1301.49</td>
<td>2.61,3965</td>
<td>&lt;.001</td>
<td>.46</td>
<td>52.34</td>
<td>(12.4)</td>
<td>60.94</td>
<td>(12.6)***</td>
<td>.69</td>
</tr>
</tbody>
</table>

* - $p \leq .05$, ** - $p \leq .01$ level, *** - $p \leq .001$.

a – Greenhouse-Geisser; b – Huynh-Feldt
scales showed more moderate but significant increases in scale score distortion as higher amounts of random responses were inserted, although four scales failed to reach a small effect size even when 90% of item responses were replaced with random responses. Of the 37 scales examined, 20 (2 Higher Order Scales, 6 Restructured Clinical Scales, and 12 Specific Problems Scales) showed mean scales scores that were distorted 5 T Score points or more in the 90% random responding condition.

Pattern analyses.

Another issue to be addressed when considering whether random responding affects the interpretation of scores on substantive scales in is whether patterns of scale elevations change as a function of random responding. As noted earlier, order of elevation plays a role in the interpretation guidelines in the test. To answer this question, \( \chi^2 \) (Chi Square) analyses were used to compare elevation patterns for the H-O Scales at each level of random response insertion with the 0% random response insertion condition (for example, 0 vs. 30%, 0 vs. 60%, and 0 vs. 90%). Results for the \( \chi^2 \) Analyses across samples can be found in Table 15. Each sample was divided into five groups represented by the data columns. The first group denotes protocols where EID was elevated and was more than 5 T-score points higher than either THD or BXD. Protocols placed in the THD group produced an elevated THD scale score that was more than 5 T-score points higher than EID and BXD. BXD was similarly sorted. WNL refers to protocols where none of the Higher-Order Scales were elevated. The multiple elevation category refers to protocols where at least one, but usually more than one scale was elevated and the
Table 15. The Effects of 30%, 60%, and 90% Random Response Insertion on the Pattern of MMPI-2 RF

Higher Order Scale Elevations

College Sample ($N = 1491$)

<table>
<thead>
<tr>
<th></th>
<th>EID</th>
<th>THD</th>
<th>BXD</th>
<th>WNL</th>
<th>Multiple Elevations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$ (%)</td>
<td>$N$ (%)</td>
<td>$N$ (%)</td>
<td>$N$ (%)</td>
<td>$N$ (%)</td>
</tr>
<tr>
<td>0%</td>
<td>145 (10%)$^a$</td>
<td>99 (7%)</td>
<td>169 (11%)</td>
<td>998 (67%)</td>
<td>80 (5%)</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30%</td>
<td>87 (6%)</td>
<td>534 (36%)</td>
<td>108 (7%)</td>
<td>579 (39%)</td>
<td>183 (12%)</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60%</td>
<td>16 (1%)</td>
<td>1178 (79%)</td>
<td>30 (2%)</td>
<td>90 (6%)</td>
<td>177 (12%)</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90%</td>
<td>0 (0%)</td>
<td>1426 (96%)</td>
<td>2 (&lt;1%)</td>
<td>1 (&lt;1%)</td>
<td>62 (4%)</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Outpatient Sample (N = 900)

<table>
<thead>
<tr>
<th></th>
<th>EID</th>
<th>THD</th>
<th>BXD</th>
<th>WNL</th>
<th>Multiple Elevations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>376 (42%)</td>
<td>64 (7%)</td>
<td>79 (9%)</td>
<td>277 (31%)</td>
<td>104 (12%)</td>
</tr>
<tr>
<td>30%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>209 (23%)</td>
<td>271 (30%)</td>
<td>43 (5%)</td>
<td>156 (17%)</td>
<td>221 (25%)</td>
</tr>
<tr>
<td>60%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>33 (4%)</td>
<td>661 (73%)</td>
<td>23 (3%)</td>
<td>26 (3%)</td>
<td>157 (17%)</td>
</tr>
<tr>
<td>90%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>2 (&lt;1%)</td>
<td>850 (94%)</td>
<td>3 (&lt;1%)</td>
<td>2 (&lt;1%)</td>
<td>43 (5%)</td>
</tr>
</tbody>
</table>

### Inpatient Sample (N =1913)

---
<table>
<thead>
<tr>
<th>EID</th>
<th>THD</th>
<th>BXD</th>
<th>WNL</th>
<th>Multiple Elevations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>0%</td>
<td>Random</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>699 (37%)</td>
<td>223 (12%)</td>
<td>167 (9%)</td>
<td>643 (34%)</td>
</tr>
<tr>
<td>30%</td>
<td>Random</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>416 (22%)</td>
<td>594 (31%)</td>
<td>122 (6%)</td>
<td>388 (20%)</td>
</tr>
<tr>
<td>60%</td>
<td>Random</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>77 (4%)</td>
<td>1379 (72%)</td>
<td>36 (2%)</td>
<td>60 (3%)</td>
</tr>
<tr>
<td>90%</td>
<td>Random</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 (&lt;1%)</td>
<td>1822 (95%)</td>
<td>5 (&lt;1%)</td>
<td>5 (&lt;1%)</td>
</tr>
</tbody>
</table>

Criminal forensic Sample (N = 995)
<table>
<thead>
<tr>
<th>EID</th>
<th>THD</th>
<th>BXD</th>
<th>WNL</th>
<th>Multiple Elevations</th>
<th>( \chi^2 )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td>153 (15%)</td>
<td>111 (11%)</td>
<td>159 (16%)</td>
<td>466 (47%)</td>
<td>106 (11%)</td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30%</td>
<td>78 (8%)</td>
<td>347 (35%)</td>
<td>104 (11%)</td>
<td>293 (29%)</td>
<td>173 (17%)</td>
<td>212.98</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60%</td>
<td>17 (2%)</td>
<td>756 (76%)</td>
<td>22 (2%)</td>
<td>42 (4%)</td>
<td>158 (16%)</td>
<td>1056.47</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90%</td>
<td>1 (&lt;1%)</td>
<td>958 (96%)</td>
<td>2 (&lt;1%)</td>
<td>2 (&lt;1%)</td>
<td>32 (3%)</td>
<td>1473.94</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Civil forensic Sample (\( N = 1521 \))

<table>
<thead>
<tr>
<th>EID</th>
<th>THD</th>
<th>BXD</th>
<th>WNL</th>
<th>Multiple Elevations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage</td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>0%</td>
<td>504 (33%)</td>
<td>81 (5%)</td>
<td>82 (5%)</td>
<td>746 (49%)</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30%</td>
<td>291 (19%)</td>
<td>451 (30%)</td>
<td>51 (3%)</td>
<td>463 (30%)</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60%</td>
<td>59 (%5)</td>
<td>1142 (75%)</td>
<td>19 (1%)</td>
<td>77 (5%)</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90%</td>
<td>3 (&lt;1%)</td>
<td>1456 (96%)</td>
<td>0 (0%)</td>
<td>2 (&lt;1%)</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* – percentages in rows may not add up to 100% due to rounding
difference between the highest scales was less than 5 T-score points.

In looking at the college sample, we can see that only 10% of students produced a protocol with an elevation on EID that was both above 65T and more than 5 T-score points higher than the next scale. Seven percent of the students produced such an elevation on THD and 11% produced such an elevation on BXD. In contrast, 67% of the sample produced a profile where none of the Higher Order Scales were above 65T. Five percent of the college students produced multiple elevations. The relation between scale elevation pattern and levels of random responding was significant ($\chi^2(4, N=1491) = 478.53, p < .001$), at the 30% random response insertion level. Fewer students were elevated on EID, and more than five times more individuals (36% of the sample) were elevated on THD at this level of random responding. The percentage of profiles that did not have elevations on any scales dropped to 39.

The relation between level of random response insertion and score pattern was significant when comparing the 0% and 60% random response insertion levels as well ($\chi^2(4, N=1491) = 1906.54, p < .001$). Here the effect is even more dramatic; with 79% of the college students now producing a clinically elevated score on THD that is at least five T-Score points higher than the scores on EID and BXD.

Finally, when comparing the zero percent random response condition to the 90% random response condition in college students, the association between level of random response insertion and profile configuration was stronger yet ($\chi^2(4, N=1491) = 2460.09, p < .001$). Here, 96% of the sample had an elevated score on THD that was more than 5 T-
score points higher than the other Higher Order scales, whereas less than one percent of the protocols had no scales with scores above 65T.

The relation between random response insertion and scale elevation pattern was significant at the 30% level in the outpatient sample ($\chi^2(4, N= 900) = 262.14, p <.001$). Fewer people were had their highest score on EID and many more people were elevated on THD in the 30% group than in the original sample. Approximately half the number of protocols in the WNL group was in the 30% random response insertion condition compared to the condition where no random responses were inserted, indicating that a large number of protocols that had originally produced an unelevated profile now showed an elevation on one or more scales. The association between level of random response insertion and relative scale elevation was significant when comparing the 0% and 60% random response insertion conditions in the outpatient condition ($\chi^2(4, N= 900) = 1028.68, p <.001$). An even more dramatic effect is seen here; one tenth the number of outpatients were elevated on EID and higher on EID than the other scales in the 60% random response insertion condition than where no random responses were inserted, and ten times more protocols produced elevated scores on THD and more than 5 T-score points higher than the other scales in the 60% random response insertion group than in the original zero percent random response insertion condition. In comparing the zero percent random response condition to the 90% random response condition, the association between level of random response insertion and highest scale elevation was significant as well ($\chi^2(4, N= 900) = 2460.09, p <.001$). Here, 94% of the sample had an
elevated score on THD that was more than 5 T-score points higher than the other Higher Order scales.

Similar patterns of results are reported in Table 15 for the inpatient, criminal forensic, and civil forensic samples. All of the $\chi^2$ analyses demonstrate that the proportion of profiles with elevated THD scores increased as the level of random response insertion increased. Across samples, at 60% random responses insertion, between three and five times more profiles had THD as the highest scale, and fewer profiles demonstrated no elevations on the Higher Order Scales. Across all samples in the 90% random response inserted condition, profiles were overwhelmingly categorized as having THD scores that were more than 5 T-score points higher than the other Higher Order Scales.
Discussion

Previous research into random responding has focused primarily on its detection. In contrast, few studies have been conducted to examine the effects of random responding on the interpretability of scale scores. This study was conducted to determine the effects of random responding on the interpretability of 37 of the MMPI-2-RF substantive scales by focusing on two separate but related questions: (1) Does prediction bias occur as a function of random responding? And (2) Does random responding distort scores on substantive scales in a manner that would affect their interpretation?

A method developed by Handel and used in previous research (Dragon, Ben-Porath, & Handel, 2012; Handel, Ben-Porath, Tellegen, & Archer, 2010) was relied upon in this investigation to randomly replace content-consistent responses with three different proportions of random responses (30%, 60%, and 90%). Random responses were inserted into archival data generated by five samples: A sample composed of individuals who are assumed to have normal psychological functioning (college students), individuals with psychological difficulties who are still functioning in a community setting (outpatients), individuals with severe psychological problems that prevented them from functioning in the community (psychiatric inpatients), individuals who are undergoing examination because of legal problems (criminal forensic evaluatees), and individuals pursuing disability claims because of physical difficulties (civil forensic
evaluatees). These samples were chosen because they exhibited different patterns of scale elevations on the Higher Order, Restructured Clinical, and Specific Problems scales of the MMPI-2-RF related to their general level of psychological functioning as well as their specific current difficulties. These patterns reflect differences in how the individuals in these samples perceive the world and how they function within it.

Correlational analyses were conducted to determine whether prediction bias occurs as a function of random responding. In all, 179 different correlations were examined. Over one-third of the associations between MMPI-2-RF scales and extra-test criteria were significantly impacted statistically, even at 30 percent random response insertion; however, the average changes in correlations in this condition were small (.05 or less) across samples. With 60 percent random responses inserted, correlations became attenuated and the majority of the comparisons demonstrated statistically and clinically significant differences from the original correlations. Most of these scale scores had correlations with extra-test criteria that were approximately half of the content-consistent correlation between the MMPI-2-RF scale and its criterion measure. At 90 percent random response insertion, the correlations were so significantly attenuated that actual significant associations between MMPI-2-RF scales and extra-test criteria disappeared, rendering the original pattern and strength of relationships with extra-test correlates undetectable. Overall, these results clearly demonstrated that although the predictive validity of the scales was minimally affected at low levels of random response insertion, it became increasingly compromised as the level of random responding increased.
This pattern is consistent with previous studies (Archer, Fontaine, & McCrae, 1998; Handel et al., 2010; Hough et al., 1990) that examined how validity coefficients attenuate as the percentage of random responding increases. This study extends previous findings in three ways. First, it illustrates how correlations with extra-test criteria become significantly attenuated as extreme amounts of random responses are inserted. Second, by only including substantial correlations to begin with, it was possible to demonstrate how significant the impact of random responding can be, even at moderate levels. Third, it illustrated that correlations between extra-test criteria and scales composed of less frequently endorsed items attenuate more rapidly than scales composed of more frequently endorsed items.

Interestingly, though no correlations were above .20 at extreme amounts of random responding, correlations with extra-test criteria appeared to be most preserved at lower levels of random responding in the inpatient group. This is likely due to the high base rate of emotional, cognitive, and behavioral problems reported by the test-takers in this sample.

Across samples, the validity of scores on EID was more robust (though still significantly affected by increasing levels of random response insertion) than the other MMPI-2-RF substantive scales. EID is the longest scale examined (41 items) and is composed, in part, of items with endorsement rates that are close to 50%. Items with endorsement rates close to 50% are less likely to be affected by random response insertion, which on average will insert 50% true responses and 50% false ones. In contrast, although THD is also one of the longer scales on the MMPI-2-RF (26 items), it
is composed of infrequently endorsed items and demonstrated rapid deflation of correlations with extra-test criteria. Thus, it is the endorsement rates of the items scored on a scale rather than its length that determines the rate at which its validity becomes affected as a function of random responding. Similarly, preservation of correlations between scales and extra-test criteria was greater in samples where more items are more frequently endorsed due to the presence of psychopathology. This is best illustrated by the inpatient sample.

The effect of random responding on predictive validity is only one concern that test users need to consider when interpreting a protocol because associations with extra-test criteria are preserved, as scale score distortion can effectively render a protocol uninterpretable. To determine whether random responding distorted scores on substantive scales in a manner that would affect their interpretation, two separate but related questions were examined: First, whether and to what extent individual scale scores are distorted under greater proportions of random responding? Second, do patterns of scale elevations change as a function of random responding?

To establish whether random responding significantly distorts scores on substantive scales in a manner that would affect their interpretation, Repeated Measures ANOVAs were conducted, followed by planned comparisons of means and standard deviations of substantive scale scores. Results indicated that for the majority of the scales across settings, scale scores became increasingly distorted as higher levels of random responses were inserted. On average, when comparing the 30% random response insertion to the content-consistent condition, effect sizes reflecting the strength of the
change of scale scores were small, with notable exceptions. Comparisons involving scales composed of lower base rate items (THD, RC6, RC8, GIC, SUI, BRF, and DSF) exhibited large effect size changes from content-consistent scale scores even at relatively low levels of random responding. Under moderate levels of random response insertion, moderate to large effects in scale scores were observed. Under extreme levels of random response insertion, large to very large effects were observed. Overall, random response insertion significantly distorted scale scores, and those distortions became more pronounced at higher levels of random response insertion, particularly for scales composed of less frequently endorsed items. Scale length did not appear to play a role in the resilience of scales to score distortion.

In examining 37 scales across 5 samples, only 14 comparisons (approximately 7%) of the analyses were not statistically significant. Some scales did not exhibit an effect in one or two samples (College sample, ACT; Outpatient sample, EID, MLS, ANP, and SAV; Inpatient sample, EID, RCd, RC3, NFC, and SAV; Forensic sample, RC3; and Civil Forensic sample, SFD, STW, and SHY); however, no scale was unaffected in three or more samples. Examination of the scale scores indicates that the effect of random response insertion was masked in samples where the content-consistent mean score was similar to the scale score mean under extreme levels of random responding. In other words, samples characterized with greater levels of psychopathology (i.e., inpatients) showed the smallest effects.

Most scales are affected by random responding, but since they are not all affected to the same degree, is the configuration of scales affected by random responding? To
determine whether individual scale score changes resulting from random responding altered their patterns of elevation, Chi Square analyses were conducted with the H-O Scales. These only focused on which scale had the highest score in each protocol across various levels of random responding. Criteria for being the highest scale required that the scale be clinically elevated (above 65T) and at least five T-score points higher than the other two H-O scales. At 30% random response insertion, all of the samples exhibited at least three times more protocols with THD as the highest elevation than they did under content-consistent responding. At 60% random response insertion, THD was the highest, most prominent scale in 70% to 80% of protocols across samples. At 90% random response insertion, 94% or more of all of the samples had scores on THD that were elevated and significantly higher the other H-O scales. Furthermore, less than one percent of any sample was composed of protocols producing scores below 65T on all H-O scales at this level of random response insertion.

As expected, the proportion of protocols that produced scores on THD that were both above 65T and five points higher than the next highest scale increased dramatically as the percentage of random responses increased. Overall, these results illustrate that a sizeable proportion of protocols would be misclassified as reflecting thought dysfunction even at lower levels of random responding. At moderate and extreme levels of random responding, the vast majority of each sample would be so misclassified.

As discussed earlier in this paper, some authors (Costa & McCrae, 1997; McGrath et al., 2010) contend that predictive validity is sufficient grounds to interpret scale scores on potentially content-inconsistent protocols. Their position presupposes that a
correlation between a test score and an extra-test criterion is necessary and sufficient
evidence of its interpretability. However, as this research demonstrates, substantive scale
scores may continue to exhibit correlations with extra-test criteria under partial random
responding conditions and yet provide misleading interpretive information due to score
distortion.

**Implications for MMPI-2-RF Interpretability**

These results demonstrate why test users need to consider the score on VRIN-r, which is designed to detect random responding on the MMPI-2-RF, before attempting to interpret substantive scale findings on the test. Failure to consult VRIN-r creates the risk of misinterpreting protocols marked by significant amounts of scale score distortion as a result of random responding. Because even lower proportions of random responding impact some scales significantly (e.g., THD, RC6, RC8, and SUI), test users should consider the potential impact that random responding might have played in elevations on these scores in protocols of borderline validity (VRIN ≥ 70T but less than 80T).

Test users should also consider what impact random responding may play on the configuration of scale scores. When a test-taker engages in random responding, patterns of scores on a protocol are likely to misrepresent his or her relative standing on the broad domains of affect, thought, and behavior. Particularly, test users should be cautious in interpreting elevations on scales such as THD, RC6, and RC8 as indicative of thought disorder if random responding is suspected, and no external evidence of thought disorder is present. In addition, scale scores on SUI may be distorted under random responding
conditions. Nonetheless, great caution must be exercised in situations where a test-taker’s life is concerned, and elevations on SUI should be investigated no matter the potential cause for elevation.

Limitations and Future Investigations

Random response insertion, as opposed to test-taker-generated random responses, was used in this investigation to determine how scales are impacted by random responding. Some authors (Costa & McCrae, 1997) have argued that random test-taker-generated responses are qualitatively different from those generated by computer, stating that test takers may respond content-consistently albeit in unusual ways. However, no data have been provided to back up this contention, and the six studies that have explored test-taker-generated random responses have not produced different results from studies using computer-generated data. Future research could address this issue more directly, comparing computer-generated to test-taker-generated random responses, and allowing researchers to use computer-generated data with greater confidence in the generalizability of their results.

Another potential limitation of this study is that random response insertion was not proportionally weighed in a way that reflects the typical pattern of test-taker random responding. For example, past research has indicated that test takers are more likely to respond randomly to items that appear at the end of the MMPI-2 booklet than to items in other parts of the test (Berry et al., 1992). Additional research should be undertaken to determine whether test-takers are more likely to respond randomly in the later part of the
MMPI-2-RF, which is considerably shorter. Future researchers may develop algorithms that better mimic test-taker-generated random responses more accurately and thus increasing the likelihood that their results reflect how test-takers actually do respond randomly.

This research addresses a larger question about the effect of non-content-based invalid responding on substantive scale score interpretability. To address this larger issue, future research should examine how other types of non-content-based invalid responding impact the interpretability of substantive scales and how well validity scales protect against interpreting protocols at different levels of random responding, as this study clearly illustrates how even partial random responding can impact both the predictive validity and interpretability of scales. Future research should also continue to explore ways to detect partial random responding, particularly concentrating on levels where traditional random responding measures appear to be less effective.

**Conclusion**

Clinicians need to know that they can trust scale scores to provide an accurate picture of a test-taker. This study illustrates how random responding impacts the interpretive validity of MMPI-2-RF scales in three ways: how scale scores are related to extra-test criteria, the ability of the scales to predict a test-taker’s actual standing on a construct of interest, and the relative prominence of problems experienced by a given test-taker. In addition, this study informs test users’ how scores on a particular scale may be biased when interpreting a protocol of borderline validity. Test users who are not
conscientious about assessing for randomly responding or users of tests that do not provide good measures of random responding run the risk of reading inaccurate conclusions about test-takers.
References


Handel, R. W., Ben-Porath, Y. S., Tellegen, A., & Archer, R. P. (2010). Psychometric functioning of the MMP-2-RF VRIN-r and TRIN-r scales with varying degrees of


