Summer 2015

Effects of Varying Degrees of Fixed and Random Responding on the Validity of Score Interpretation for the SP and PSY-5 Scales of the MMPI-2-RF

Joseph Brooks Minifie
Old Dominion University

EFFECTS OF VARYING DEGREES OF FIXED AND RANDOM RESPONDING ON THE VALIDITY OF SCORE INTERPRETATION FOR THE SP AND PSY-5 SCALES OF THE MMPI-2-RF

by

Joseph Brooks Minifie
B.S. May 2008, Trinity College
M.A. July 2013, Norfolk State University

A Dissertation Submitted to the Faculties of Eastern Virginia Medical School,
Norfolk State University and Old Dominion University
in Partial Fulfillment of the Requirements of

DOCTOR OF PSYCHOLOGY

CLINICAL PSYCHOLOGY

VIRGINIA CONSORTIUM PROGRAM IN CLINICAL PSYCHOLOGY
August 2015

Approved by:

_____
Richard Handel, Ph.D. (Chair)
Eastern Virginia Medical School

_____
Robert Archer, Ph.D. (Member)
Eastern Virginia Medical School

_____
Desideria Hacker, Ph.D. (Member)
Norfolk State University

_____
Serina Neumann, Ph.D. (Member)
Eastern Virginia Medical School

_____
James Paulson, Ph.D. (Member)
Old Dominion University

ABSTRACT

EFFECTS OF VARYING DEGREES OF FIXED AND RANDOM RESPONDING ON THE VALIDITY OF SCORE INTERPRETATION FOR SP AND PSY-5 SCALESOF THE MMPI2-2-RF

Joseph Brooks Minifie
Virginia Consortium Program, 2015
Director: Dr. Richard Handel

The Minnesota Multiphasic Personality Inventory – 2 Restructured Form (MMPI-2-RF; Tellegen & Ben-Porath, 2008/2011) is a widely used self-report measure of psychopathology and personality. However, the self-report format of the MMPI-2-RF suggests that interpretation of its scales and the clinical recommendations that follow are vulnerable to invalid response styles. This dissertation builds upon previous research (Handel, Ben-Porath, Tellegen, & Archer, 2010) to examine the effect of random and fixed responding, as measured by the VRIN-r and TRIN-r Scales, on the 28 SP and PSY-5 Scales. A computer simulation procedure was used to insert increasing degrees of inconsistent responding into protocols from two large samples (N = 2, 276 and N = 704). Results indicated that increasing degrees of inconsistent responding increase SP and PSY-5 Scale mean T-scores and weaken external criterion validity. Further, certain SP and PSY-5 Scales evidenced large changes in mean T-scores at relatively low levels of simulated inconsistent responding. Implications of these results and future areas of investigation are discussed.

ACKNOWLEDGMENTS

Of those who have helped, supported, and loved me, a few deserve special recognition here. First, a very special thanks to my dissertation advisor, Dr. Richard Handel, whose diligence, support, and belief in me have been vital elements for this project and my broader personal and professional growth. A special thanks to the members of my committee; your patience, guidance, and flexibility have been most appreciated. I would also like to acknowledge Dr. Danielle Burchett, who graciously granted me access to her dissertation. Second, my graduate school peers, particularly L.H., M.B., J.K., J.S.P., and L.W. deserve special recognition. Your support and humor have kept me grounded and smiling in the most uncertain of times. Third, members of my family, most importantly my wife, Megan, and my mother, brother, and late father have and will continue to be the people I have worked so hard for these past years. And last, but certainly not least, a special thanks to our two dogs, Scout and Bella, who have taught me the invaluable lesson that good things come to those who wait.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

Table Page

CHAPTER I

INTRODUCTION

The Minnesota Multiphasic Personality Inventory – 2 – Restructured Form (MMPI-2-RF; Tellegen & Ben-Porath, 2008/2011) represents the newest development of one of the most frequently used psychological assessments (Camara, Nathan, & Puente, 2000). Its predecessor, the MMPI-2 (Butcher, Graham, Ben-Porath, Tellegen, Dahlstrom, & Kaemmer, 2001), contains 567 items that comprise 10 Validity Scales, 10 Clinical Scales, and 93 additional scales and subscales. These scales and subscales, based on a large normative sample, provide detailed information about protocol validity, psychopathology, and personality. Furthermore, a wide research base exists for the MMPI-2 (Graham, 2012).

Despite these strengths, however, a series of criticisms were lodged against the MMPI-2 (Ben-Porath, 2012). Chief among these criticisms was that the MMPI-2 Clinical Scales, designed to be a primary source of information about psychopathology and personality, no longer satisfied modern psychometric standards. Tellegen, Ben-Porath, McNulty, Arbisi, Graham, and Kaemmer (2003) addressed these weaknesses by publishing the Restructured Clinical (RC) Scales, which are currently used in conjunction with the Clinical Scales in MMPI-2 protocol interpretation. However, the publication of the RC Scales proved to be the first step in a process that ultimately resulted in the MMPI-2-RF, as similar scale development methods were used to explore how the MMPI-2 item pool could be used to develop entirely new scales.

As compared to the MMPI-2, the MMPI-2-RF contains 338 items that comprise 10 Validity Scales, three Higher-Order Scales, nine RC Scales, and 30 additional scales.

Major advantages of the MMPI-2-RF include decreased administration time, improved psychometric functioning, and less complex protocol interpretation (Ben-Porath, 2012; Graham, 2012). Furthermore, the addition of the Higher-Order Scales, retention of the RC Scales, and addition of Specific Problem (SP) Scales afford the examiner a hierarchical protocol interpretation strategy. As a relatively new measure, additional research studies on certain areas of MMPI-2-RF psychometrics are needed. Therefore, the overall purpose of this dissertation will be to add to the growing MMPI-2-RF literature base. Specifically, an examination of how interpretation of certain substantive scales may change under conditions of invalid responding is proposed.

CHAPTER II

REVIEW OF THE LITERATURE

The MMPI-2-RF is shorter than its counterpart, the MMPI-2. In 338 items, researchers and clinicians are presented with a broadband assessment of psychopathology and personality (Ben-Porath, 2012). Yet, as with all self-report measures, the presence of invalid responding can significantly distort scores on the MMPI-2-RF, leading to inaccurate interpretations and recommendations. This literature review will focus on the assessment of threats to protocol validity, specifically variable and fixed inconsistent responding. It will also focus on the SP and Personality Psychopathology Five (PSY-5) Scales of the MMPI-2-RF. In reviewing the development and empirical support of these substantive scales, the richness of the clinical information they provide will be contrasted with their susceptibility to patterns of invalid responding. This susceptibility, in turn, will be proposed as the focus of this dissertation.

Non-Content-Based Invalid Responding

In developing the original MMPI, Hathaway and McKinley (1943) were aware of the potential for psychiatric and medical patients to distort responses on self-report measures (Graham, 2006; Hoelzle, Nelson, & Arbisi, 2012). In the decades since the publication of the MMPI, increasingly sophisticated and accurate measures of response distortion have been developed. These developments, in turn, have led to improvements in the examiner's ability to discriminate between psychological functioning assessed by the measure and inconsistent responding (Tellegen, 1988), an essential piece of information when evaluating protocol validity.

Accordingly, assessing the validity of a test taker's responses is a necessary first step in protocol interpretation (Ben-Porath, 2012). Threats to protocol validity occur in two primary forms. First, content-based invalid responding (CBIR) concerns test takers who can read and comprehend items, yet respond to item content in a skewed manner that either amplifies or minimizes psychopathology. Furthermore, test takers can engage in CBIR intentionally or unintentionally.

Alternatively, non-content-based invalid responding (NCBIR) occurs when test takers respond to items without regard for item content or in a manner that prohibits item scoring (e.g., leaving an item blank; Ben-Porath, 2012). Some causes of NCBIR include poor reading ability, defensiveness, uncooperativeness, and limited insight. Like CBIR, NCBIR can also be intentional or unintentional. Both forms of invalid responding present significant concerns for protocol interpretation, as the validity of substantive score interpretation decreases as invalid responding increases.

The assessment of NCBIR occurs in three forms (Ben-Porath, 2012). First, nonresponding refers to a response style marked by leaving items blank or with both response options filled out. Second, random responding refers to the tendency of a test taker to fill out items in a random manner without regard for item content. While test takers can respond randomly to entire protocols, it is more likely that random responding occurs in varying degrees. Third, fixed responding concerns test takers who respond in a systematic manner to items without regard for item content (e.g., responding *true* to five items or *false* to five items, etc.). As with random responding, fixed responding can also occur in varying degrees.

The MMPI family of assessments has substantial empirical support. Yet, research examining NCBIR, specifically the measures of variable and fixed inconsistent responding on the MMPI-2-RF, is relatively limited. In the following sections, literature concerning the development, empirical support, and importance of these measures will be reviewed.

Validity Scales on the original MMPI. Three validity scales were published with the release of the original MMPI (Hathaway & McKinley, 1943; Ben-Porath, 2012); one more was added several years later. First, Cannot Say (CNS) assessed the raw number of responses the test taker could not respond to with either *true* or *false*. High CNS scores were considered problematic because missing responses artificially lowered other MMPI scales. Second, the Lie (L) score measured the extent to which a test taker was attempting to create a positive and socially acceptable image. Third, the Infrequency (F) Scale was intended to assess random responding and both unintentional and intentional overreporting. Finally, the K Scale (Meehl & Hathaway, 1946), a measure of underreporting, was added in 1946. Thus, the F Scale was the only measure of random responding available on the original MMPI, and assessment of fixed responding was not possible.

A high score on F indicated that the test taker had endorsed a large amount of items that were infrequently endorsed in the MMPI normative sample (Buechley & Ball, 1952). The F scale was composed of items that were endorsed by 10% or less of the participants in the MMPI normative sample. While a high F score was suggestive of random responding, it was not possible to discriminate between a score indicative of random responding and one indicative of overreporting or actual, serious

psychopathology (e.g., schizophrenia). Furthermore, items on F were found within the first 300 of the 566 total MMPI items; therefore, an assessment of invalid responding on the second part of the MMPI was not possible. In response to these concerns, research into how other MMPI items could be used to create new validity scales began soon after the MMPI was released (Hoelzle et al., 2012).

One of the first of these additions was Buechley and Ball's (1952) Test-Retest (TR) Scale. The TR Scale consisted of 16 identical item pairs distributed throughout the MMPI item pool. The number of item pairs answered inconsistently, then, was a measure of inconsistent responding. Furthermore, the TR Scale could be used in conjunction with F to determine whether high F scores were indicative of random responding, overreporting, or actual pathology. Difficulties in determining an appropriate TR Scale cutoff score, however, surfaced in the literature. While Buechley and Ball recommended a cutoff score of three or more, Greene (1979) recommended a score of four or more and Nichols, Greene, and Schmolck (1989) recommended a score of greater than six. Furthermore, the TR Scale was ineffective in detecting patterns of fixed responding. To address these shortcomings Greene (1978) developed the Carelessness Scale, which consisted of pairs of items with identical content and items representing "psychological opposites." The Carelessness Scale added significantly to the TR Scale in identifying invalid protocols. Additionally, it improved detection of fixed responding (Nichols et al., 1989).

Yet as the number of new validity scales increased, so too did the number of interpretive guidelines for these scales. For example, Rogers, Dolmetsch, and Cavanaugh (1983) recommended an F Scale score greater than 80 and TR Scale score

greater than 4 as a basis to discriminate random versus nonrandom responders. Nichols et al. (1989) developed a series of six decision rules designed to improve classification accuracy. While their results demonstrated improved performance over traditional validity scales, the proposed decision rules were both lengthy and complex.

Thus, assessment of random and fixed responding in the MMPI era had two significant shortcomings. First, increasing the accuracy of detecting protocols invalid due to random responding came with lengthy and complex scoring procedures. Second, scales designed to assess fixed responding were generally less developed than those intended to assess random responding. These concerns were addressed with the first major revision of the MMPI.

Inconsistent responding and the MMPI-2. With the release of the MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) three new validity scales were added: Variable Response Inconsistency (VRIN), True Response Inconsistency (TRIN), and Back F (F_B; Ben-Porath, 2012). These scales were initially viewed as supplementary to CNS, L, F, and K, and as an experimental addition to the MMPI-2. It was not until the second edition of the MMPI-2 manual (Butcher et al., 2001) that VRIN, TRIN, and additional validity scales were added to the validity scale profile (Ben-Porath, 2012).

Prior to their addition to the MMPI-2, VRIN and TRIN scales were originally created for the Multidimensional Personality Questionnaire (MPQ; Tellegen, 1995/2003). In designing these scales, Tellegen dichotomized inconsistent responding as either (a) responses filled out in the same manner (e.g., answering all *true*) or (b) responses filled out in the opposite manner (e.g., answering *true* and *false* randomly;

Tellegen, 1988, as cited in Ben-Porath, 2012). Therefore, in creating the MPQ version of VRIN, Tellegen created item pairs with similar content. Conversely, TRIN was composed of pairs of items containing dissimilar content. Tellegen reported results showing the sensitivity of each scale in detecting their respective type of inconsistent responding. Furthermore, he found that both scales were necessary to evaluate inconsistent responding, as "neither scale was effective at detecting the type of inconsistent responding detected by the other" (Ben-Porath, 2012, p. 151).

In adapting the MPQ versions of VRIN and TRIN to the MMPI-2, Butcher et al. (1989) intended to use these scales in the same manner (Ben-Porath, 2012). However, these scales were modified in three main ways. First, Tellegen had sought to reduce the effects of individual personality traits on VRIN and TRIN by selecting items from multiple content areas. Similarly, Butcher and colleagues selected VRIN and TRIN item pairs that were minimally correlated with Clinical Scale scores to reduce the effect of psychopathology on VRIN and TRIN. Second, item pairs were added to VRIN that were also found in TRIN; VRIN now contained item pairs with both similar and opposite content. While this change increased the degree of overlap between these two scales, it also broadened the ability of VRIN to detect inconsistent responding. Finally, Butcher and colleagues realized that inconsistent responses could be asymmetrical: responding false to both items in a VRIN pair, for example, would be evidence of inconsistency, whereas responding true to both items would not. Therefore, item pair correlations and observed to expected frequencies were calculated to arrive at the final VRIN and TRIN scales.

The result of these modifications produced a VRIN Scale consisting of 67 item pairs and a TRIN Scale of 23 (Ben-Porath, 2012). Given their role as experimental scales in the 1989 edition of the MMPI-2, interpretive guidelines were somewhat unclear. However, the recommended T-score of 80 received external empirical support for VRIN (Archer, Fontaine, & McCrae, 1998; Berry et al., 1991; Paolo & Ryan, 1992) and TRIN (Handel, Arnau, Archer, & Dandy, 2006). As evidence for the use of these scales increased, VRIN and TRIN were moved to the forefront of profile interpretation with the revision of the MMPI-2 manual (Butcher et al., 2001). Only CNS was placed before these scales.

Empirical support for VRIN and TRIN. The addition of VRIN, TRIN, and F_B (which also assessed overreporting) represented an improvement in NCBIR identification as compared to the MMPI (Sprock, 2000). The sensitivity of VRIN to partially random (Archer et al., 1998; Berry et al., 1991; Berry et al., 1992) and entirely random (Paolo & Ryan, 1992; Pinsoneault, 2007; Sewell & Rogers, 1994; Wetter, Baer, Berry, Smith, & Larsen, 1992) MMPI-2 protocols was demonstrated through a series of experiments using computer-generated and participant-generated random and valid protocols. Archer et al. (1998), Berry et al. (1991), and Paolo & Ryan (1992) provided general support for the Butcher et al. (1989) recommendation that a T score of 80 was indicative of an invalid protocol, although Berry et al. (1991) raised concerns about the sensitivity and specificity of VRIN under different base rates of random responding. Furthermore, even under conditions where a protocol was determined to be invalid, Archer et al. (1998) reported that concurrent validity of substantive scales remained relatively strong.

Despite this strong empirical support, several criticisms emerged against VRIN. First, as discussed above, concerns were raised about the sensitivity and specificity of VRIN under varying base rates of inconsistent responding (Berry et al., 1991; Gallen & Berry, 1996). These concerns prompted research into additional validity indexes, including: (a) |F-Fb|, VRIN+|F-Fb|, and F+Fb+|F-Fb| (Cramer, 1995; Gallen & Berry, 1996; Greene, 1991), (b) the use of VRIN confidence intervals (Charter & Lopez, 2003), and (c) VRIN subscales derived by dividing the MMPI-2 protocol into segments (Pinsoneault, 2007). In a process similar to MMPI additions, while these experimental scales often resulted in improved sensitivity, their complex scale calculations posed challenges to both researchers and clinicians.

Second, concerns also arose about VRIN's decreased sensitivity to partially random protocols. Clark, Gironda, and Young (2003) reported that VRIN evidenced decreased sensitivity to random responding present in the second half of the MMPI-2 protocol. Cramer (1995) reported that VRIN could not discriminate between all degrees of profile randomness, and Pinsoneault (2007) reported data suggesting that VRIN exhibited poorer performance with protocols that were less than 100% random. Third, and finally, hand-scoring the VRIN was a difficult process that often resulted in errors (Iverson & Barton, 1999), leading Sewell and Rogers (1994) to develop a 16-item screening measure to determine if VRIN even needed to be scored at all.

In comparison to VRIN, the available research on TRIN is limited (Handel et al., 2006; Hoelzle et al., 2012). However, the existing literature supports the use of TRIN. Using the MMPI-2 and MMPI-A normative sample protocols, Handel and colleagues randomly introduced increasing levels of both *true* and *false* responses. Their results

revealed that TRIN demonstrated strong sensitivity to both acquiescent (i.e., responding *true*) and counter-acquiescent (i.e., responding *false*) fixed responding.

Validity scales on the MMPI-2-RF. The 338-item MMPI-2-RF contains a comprehensive revision of all existing scales and the addition of new scales (Ben-Porath, 2012). Revisions to VRIN (relabeled VRIN-r), TRIN (relabeled TRIN-r), and additional validity scales occurred last. Using the framework provided by the MMPI-2 versions of VRIN and TRIN, Tellegen and Ben-Porath (2008/2011, as cited in Ben-Porath, 2012) sought to achieve two primary goals. First, they sought to remove item overlap between VRIN and TRIN, given that the MMPI-2 versions of these scales shared 10 items. Second, because the RF was considerably shorter than the MMPI-2, they "used a different approach to reducing the impact of substantive content variance on scores on these scales" (Ben-Porath, 2012, p. 160).

Tellegen and Ben-Porath (2008/2011) addressed the first goal by restricting potential items for VRIN-r to item pairs that were positively correlated with each other (Ben-Porath, 2012). In other words, VRIN-r item pairs were selected so that test takers would provide the same response (e.g., responding *true* to both items) under valid conditions. Under conditions of random responding, therefore, the test taker would response *true* to one item and *false* to another. Conversely, potential TRIN-r pairs were selected if item pairs were negatively correlated with each other.

Regarding the second goal, VRIN-r and TRIN-r item pairs had to satisfy five criteria (Ben-Porath, 2012). First, item pairs had to evidence the aforementioned patterns of correlations in two clinical samples. Second, the chance of answering both items in a pair (i.e., the observed frequency) indicative of inconsistent responding had to be lower

than the frequency expected by chance. Third, responses to item pairs that would be counted towards VRIN-r and TRIN-r raw scores were judged by Tellegen and Ben-Porath (2008/2011) to contain inconsistent content. Fourth, correlations were calculated between item pairs keyed to indicate inconsistent responding and the same item pairs keyed to indicate consistent (i.e., valid) responding (Ben-Porath, 2012; Handel et al., 2010). Item pairs that evidenced low correlations in these analyses were eligible for inclusion because they were considered to have minimal associations with actual RF content. In other words, item pairs on VRIN-r and TRIN-r, now drawn from a smaller item pool, would not be affected by psychopathology. Fifth, and finally, individual items in an item pair keyed to indicate inconsistency (e.g., by two *true* responses) could not be used in another item pair in which both items were also keyed by two *true* responses. For example, if a *true* response to the hypothetical item "I am fat" and a *true* response to the hypothetical item "I am thin" indicated acquiescent responding, then a *true* response to "I am fat" could not be used with a *true* response to "I am skinny" to also indicate acquiescent responding. In essence, this criterion minimized the effect of individual items on VRIN-r and TRIN-r raw scores.

The result of these selection criteria was the MMPI-2-RF scales VRIN-r and TRIN-r (Ben-Porath, 2012). VRIN-r consists of 53 item pairs; TRIN-r of 26. Interpretative recommendations for these scales remain the same as those used in the MMPI-2: T-scores of greater than or equal to 80 on VRIN-r and TRIN-r indicate an invalid protocol due to random or fixed responding, respectively (Tellegen & Ben-Porath, 2008/2011; Handel et al., 2010). Although each scale contains a similar number of items when

compared to the MMPI-2 versions, the RF versions were designed to assess inconsistent responding across a greater percentage of available items.

Psychometric functioning of VRIN-r and TRIN-r. Specific reliability estimates, including test-retest, internal consistency, and standard errors of measurement (SEM), for VRIN-r and TRIN-r are presented in the MMPI-2-RF *Technical Manual* (Tellegen & Ben-Porath, 2008/2011, as cited in Ben-Porath, 2012). Generally, however, VRIN-r and TRIN-r evidenced fairly low estimates of reliability. Ben-Porath (2012) acknowledged these lower estimates, stating that some attenuation of reliability is expected given that "the Inconsistency Scales were designed to be content-free indicators and have an even greater restriction of range than do the other validity indicators" (p. 165). Ben-Porath also discussed the relatively higher SEMs of the Validity Scales, including VRIN-r and TRIN-r. He stated that as compared to the Substantive Scales, greater deviations from the norm are necessary "to raise substantial concerns about the validity of a test protocol" (Ben-Porath, 2012, p. 165-166).

Validity information about the inconsistency scales exists in three primary forms: (a) internal correlate data examining the relationship between the MMPI-2 Inconsistency Scales and VRIN-r and TRIN-r (Handel, Ben-Porath, Tellegen, & Archer, 2007, as cited in Ben-Porath, 2012), (b) one external empirical study using simulated random and fixed responding (Handel et al., 2010), and (c) one unpublished dissertation using simulated random responding (Dragon, 2012). With respect to the internal correlate data, Handel et al. (2007) randomly replaced responses from normative sample data and calculated correlations between MMPI-2 and MMPI-2-RF versions of VRIN

and TRIN and VRIN-r and TRIN-r, respectively. Results indicated high correlations between both versions of these scales.

External research examining the validity of VRIN-r and TRIN-r is limited; only two studies currently exist. In the first study, Handel et al. (2010) introduced "increasing levels of simulated random responding, acquiescence, and counter-acquiescence" (p. 90) using computer simulations into protocols from the MMPI-2 normative sample and into protocols from a sample of mental health inpatients. Results supported Ben-Porath and Tellegen's (2008/2011) interpretive recommendations that T-scores on VRIN-r and TRIN-r greater than or equal to 80 were indicative of invalid random or fixed responding, respectively. Furthermore, Handel and colleagues examined the effects of varying degrees of random and fixed responding on RC Scale interpretation, convergent validity, and discriminant validity. When simulated random responses were inserted into 30% of items within the protocols (corresponding to a mean VRIN-r T-score of 70), mean scale scores for RC1 (Somatic Complaints) and RC8 (Aberrant Experiences) increased by 10 T-score points. RC6 (Ideas of Persecution) mean scores were found to have increased by 15 T-score points. When acquiescent responses were inserted into 20% of protocol items (corresponding to a TRIN-r True T-score of 70), RC6 and RC8 Scale mean scores exhibited significant score distortions. Similarly, mean scores on RC1 and RC2 (Low Positive Emotions) exhibited significant score distortions when counter-acquiescent response insertion reached 30% (corresponding to a TRIN-r False T-score of 70).

Regarding the effects of simulated responding on convergent validity, Handel et al. (2010) calculated convergent and discriminant validity coefficients between the RC

Scales and an external measure. To examine how validity coefficients degraded as a function of random and fixed response insertion, Pearson *r* values for baseline (i.e., 0% response insertion) and experimental (e.g., 40% random response insertion) conditions were squared and subtracted from each other. This allowed the researchers to quantify the percentage of variance lost as a result of simulated response insertion. At random and fixed response insertion rates of 10% and 20%, convergent validity coefficients did not evidence degradation (i.e., a substantial loss of variance account for). Convergent validity coefficients for three of the nine RC Scales degraded at 30% simulated random responding, as indicated by a 4-5% loss of variance accounted for in these scales. When acquiescent response insertion was increased to 30%, validity coefficients for two of the nine scales evidenced variance accounted for reductions of 5%. When counter-acquiescent response insertion was increased to 30%, variance accounted for reductions of 3-4% were noted for three of the nine scales. At higher rates of random and fixed response insertion (e.g., 70%), variance accounted for reductions ranged from 3-11%. It should be noted that Handel and colleagues found that increasing degrees of random and fixed responding had no meaningful effects on discriminant validity coefficients.

In the second study, Dragon (2012) used the same computer simulation as Handel et al. (2010) to introduce increasing degrees of random responding, as represented by VRIN-r, into protocols from five different large samples. These samples included a comparison college sample (*N* = 1,464), an inpatient sample (*N* = 1,913), an outpatient sample (*N* = 900), a clinical forensic sample (*N* = 995), and a civil forensic sample (*N* = 1,521). Using 30, 60, and 90% rates of simulated random response insertion, Dragon examined the effects on T-scores and convergent validity coefficients for the 35 H-O,

RC, and SP Scales of the MMPI-2-RF. Convergent validity coefficients were calculated between most of these scales and a host of extra-test measures. Results of this study were highly similar to those reported Handel and colleagues. Specifically, at 30% simulated random insertion across the five experimental samples, most of the 35 H-O, RC, and SP Scales evidenced statistically significant mean T-score changes from baseline (Dragon, 2012). Further, the magnitude of change increased as simulated random response insertion rose to 60 and 90%.

Regarding the convergent validity analyses, Dragon's (2012) results indicated that increasing degrees of random responding resulted in increasing degrees of coefficient degradation between MMPI-2-RF Scales and relevant external measures. It should be noted that the greatest degradations were observed for random response rates of 60 and 90%; convergent validity coefficients were relatively robust to 30% random response insertion.

Handel et al. (2010) demonstrated the deleterious effects of random and fixed responding on RC Scale interpretation, while Dragon (2012) extended this research by demonstrating the deleterious effects of random responding on the interpretation of H-O, RC, and SP Scales. These authors and others (e.g., Burchett, 2012; Burchett & Ben-Porath, 2010) have called for an examination of the effects of random and fixed responding on the Specific Problem (SP) and the Personality Psychopathology Five (PSY-5) Scales. Accordingly, the goal of the present study is to examine the effects of simulated inconsistent responding on the Specific Problem (SP) and the Personality Psychopathology Five (PSY-5) Scales.

The Specific Problem Scales

In redesigning the Clinical Scales, Tellegen et al. (2003) published the RC Scales for use with the MMPI-2 and eventually with the MMPI-2-RF. These scales were revised primarily "to assess a major distinctive core component of an original Clinical Scale" (Ben-Porath, 2012, p. 97), in addition to reducing intercorrelations between and content heterogeneity within the Clinical Scales themselves. Thus, in order to augment the RC Scales, create a broadband assessment of psychopathology and personality, and utilize the clinical richness represented by the items contained in the MMPI-2, Ben-Porath and Tellegen (2008/2011) added and revised three primary sets of scales when creating the MMPI-2-RF.

The first two of these sets, the Higher-Order (H-O) Scales and the SP Scales, relate to the assessment of psychopathology; the third, the PSY-5 Scales, relates to the dimensional assessment of personality (Ben-Porath, 2012). The H-O Scales were designed to capture three broadband dimensions of psychopathology, labeled Emotional/Internalizing Dysfunction (EID), Thought Dysfunction (THD), and Behavioral/Externalizing Dysfunction (BXD). Conversely, the SP Scales were designed to be more narrow assessments of psychopathology. Their purpose was to (a) measure constructs not covered by the RC Scales, (b) capture facets of RC Scales needing distinct measurement (e.g., suicide), and (c) assess constructs that could be measured with MMPI-2 items that were not included in the RC Scales. A discussion of RC Scale development will augment and understanding of the SP Scales.

RC Scale development. The publication of the RC Scales in 2003 represented the first of several major scale development efforts that ultimately resulted in the MMPI-2-

RF (Ben-Porath, 2012). Furthermore, the methods used by Tellegen et al. (2003) to revise these scales were used in generating the SP Scales, to be discussed later. As discussed briefly above, two primary concerns motivated the developers of the RC Scales (Ben-Porath, 2012). First, excessive intercorrelations among the Clinical Scales had been identified. As an extreme example, a correlation of .90 was found in certain clinical samples for Clinical Scale 7 (Psychasthenia) and Clinical Scale 8 (Schizophrenia). One source of this problem was the considerable degree of item overlap between scales. This was intentional on the part of the original MMPI developers, who were trying to assess separate syndromes that shared clinical features. However, the presence of overlapping items increased measurement error and decreased discriminant validity. Another source of this problem, to be discussed later, was the presence of demoralization, a factor found in each of the Clinical Scales. Second, Clinical Scales assessed heterogeneous content. Like the first problem, this shortcoming was also by design; MMPI developers wanted to assess syndromes with several facets. As a result of this design, however, the Clinical Scales evidenced relatively weak convergent validity. Furthermore, moderately elevated scale scores were difficult to interpret.

Prior to the development of the RC Scales, attempts to address these shortcomings involved either the use of code-types or interpretation of the Harris-Lingoes Subscales, or both (Ben-Porath, 2012). Code-types involved combining two or three clinically elevated Clinical Scales to arrive at a more nuanced interpretation, while the Harris-Lingoes Subscales, a set of 31 scales created to assess the heterogeneous content within the Clinical Scales (Graham, 2012), could be used if a code-type could not be discerned.

These and other methods were not without their own weaknesses, namely relatively poor psychometric properties and increased complexity with respect to interpretation (Ben-Porath, 2012). In their revisions, then, Tellegen et al. (2003) sought to improve both efficiency and the psychometric properties of the Clinical Scales.

RC Scale development occurred in four steps (Ben-Porath, 2012). First, Tellegen, among other researchers, had recognized that in addition to assessing specific clinical syndromes, the MMPI and MMPI-2 also assessed "a broad, affectively colored construct" (Ben-Porath, 2012, p. 47) termed "demoralization." In considering that the original MMPI was constructed by comparing the keyed responses of hospital inpatients with their physically and mentally healthy friends and relatives, it becomes clear why the MMPI-2 items might have captured this general factor. However, as with item overlap, the presence of demoralization increased intercorrelations between Clinical Scales, thereby reducing discriminant validity. Thus, Tellegen identified MMPI items via factor analysis to create a measure of demoralization, which was marked by "features such as unhappiness, a poor self-concept, a sense of being overwhelmed, and a desire to give up" (Ben-Porath, 2012, p. 48). The construct represented by this measure would be used in subsequent steps to create a standalone measure of demoralization and to reduce the saturation of demoralization in other scales.

Second, Tellegen conducted factor analyses on each of the Clinical Scales to determine the number of distinctive components represented within these heterogeneous scales (Ben-Porath, 2012). These analyses were also performed in conjunction with the measure of demoralization previously discussed to determine how many items from each scale primarily loaded on this general construct. The results of the factor analyses

revealed two to four distinctive components for each scale, which included a factor representing demoralization. In addition to demoralization, several components were found to be highly similar across scales. Therefore, in choosing the components to represent each revised scale, Tellegen did not always select the largest factor identified in each scale analysis. Rather, to avoid redundancy, he selected factors for each scale that were not going to be represented by other scales.

The third step involved a series of analyses designed to produce internally consistent and distinctive Seed Scales (Ben-Porath, 2012). Items were retained if they evidenced sufficient loadings on their designated factor; items were dropped if they evidenced sufficient loadings for more than one factor. Similarly, items were retained if they were sufficiently correlated with their designated factor and they were deleted if they were more highly correlated with the other Seeds. The result of this step was the development of a set of Seed Scales used in the final step of RC Scale development.

In the final step, MMPI-2 items were recruited to create the RC Scales in a series of four sets of analyses (Ben-Porath, 2012). First, items were considered for RC Scale inclusion if they evidenced high correlations with a specific Seed Scale and low correlations with other Scales. Second, Tellegen et al. (2003) determined that the item content of provisional scales for RC7 (Negative Emotionality) and RC9 (Hypomanic Activation) was too heterogeneous (Ben-Porath, 2012). Therefore, a subset of items was identified that "unequivocally represented the intended constructs" (Ben-Porath, 2012, p. 51); items uncorrelated with this subset were removed. Third, internal consistency of each provisional scale was improved by examining the items within each provisional scale. Finally, items in each of the provisional scales were changed if these

modifications improved correlations with external criterion measures. This process resulted in a total of nine RC Scales: RCd (Demoralization), RC1 (Somatic Complaints), RC2 (Low Positive Emotions), RC3 (Cynicism), RC4 (Antisocial Behavior), RC6 (Ideas of Persecution), RC7 (Dysfunctional Negative Emotions), RC8 (Aberrant Experiences), and RC9 (Hypomanic Experiences).

Initial psychometric analyses completed by Tellegen et al. (2003) revealed that as compared to the Clinical Scales, the RC Scales evidenced "comparable to improved reliability, substantial reduction in saturation with demoralization, reduced intercorrelations, comparable to improved convergent validity, and improved discriminant validity (Ben-Porath, 2012, p. 52). Subsequent empirical investigations using a variety of mental and physical health samples have generally supported these findings. Criticisms of the RC Scales have focused on differences between these scales and the Clinical Scales. Specifically, critics have voiced concerns about how the constructs assessed by the RC Scales are not as broad as the original scales. As just reviewed, this was an intentional change that has resulted in compelling supportive evidence.

SP Scale development. In developing the SP Scales Ben-Porath and Tellegen (2008/2011) targeted several constructs (Ben-Porath, 2012). First, in analyzing the factor structure of the Clinical Scales in preparation for RC Scale development, several Clinical Scales were found to contain multiple distinctive components. As previously discussed, only one of these components was selected for RC Scale development; the other components were targeted for further development as a potential SP Scale. Second, while the RC Scales represented an improvement with respect to heterogeneity

of content as compared to the Clinical Scales, some of the RC scales still contained several facets. For example, RC1 (Somatic Complaints) was designed to assess a wide variety of somatic symptoms, such as pain and gastrointestinal complaints. Ben-Porath and Tellegen focused on these and other facets as potentially valuable sources of additional clinical information worthy of further development. Finally, constructs of particular clinical relevance (e.g., suicidal ideation and attempts) that were contained within the pool of MMPI-2 items but not in the RC Scales themselves were also targeted for possible scale development.

SP Scale development followed a multistep process similar to that used in RC Scale development (Ben-Porath, 2012). First, sets of items were created to represent each targeted construct. Generally, items were eliminated if they evidenced "excessive loadings" on the construct of demoralization. The exception to this step was, for example, the SP Scale measuring suicidal ideation, which is related to demoralization. Second, items from each list were removed if they exhibited high correlations with other item sets. This resulted in Seed Scales. Third, Seed Scales were correlated with the MMPI-2 item pool. Items were included in a scale if they (a) exhibited high correlations with a particular scale, (b) exhibited low correlations with other scales, and (c) were conceptually related to the construct targeted by the scale. Fourth, reliability was improved by removing items found to decrease scale internal consistency. Finally, and in anticipation of later validation analyses, scales were retained only if there existed scale-specific empirical correlates. These steps, in addition to subsequent analyses and feedback gathered from other experts, resulted in the 23 SP Scales. While some of the

SP Scales relate to RC Scale facets, they are unique in that they assess items not found on RC Scales; thus, they can be interpreted separately from the RC Scales.

SP Scale domains, interpretation, and psychometrics. Four domains organize the SP Scales: Somatic/Cognitive, Internalizing, Externalizing, and Interpersonal (Ben-Porath, 2012). Each scale within these domains will be reviewed with respect to information summarized by Ben-Porath and the *MMPI-2-RF Technical Manual* (Tellegen & Ben-Porath, 2008/2011), namely: the construct assessed, connections with MMPI-2 scales, and external correlates. Generally, the SP Scales demonstrate adequate reliability. Information concerning external correlates, as summarized below for each scale, was gathered from samples of community outpatients, psychiatric inpatients, forensic disability claimants, VA outpatients, and individuals seeking substance abuse treatment. Although sparse, available research not summarized by Ben-Porath will also be reviewed.

The Somatic/Cognitive Scales. Five scales are included in this domain, all of which relate to facets assessed by RC1 (Ben-Porath, 2012; Forbey, Lee, & Handel, 2010). Malaise (MLS), the first scale, was designed to assess the test taker's sense of being in poor physical health (Ben-Porath, 2012; Graham, 2012). It is conceptually similar to Hy3, the Harris-Lingoes Lassitude/Malaise subscale of Clinical Scale 3. Research conducted with Hy3 revealed that scores on this subscale were the best predictors of employee back injuries capable of interfering with occupational functioning. More recent research has also found a relationship between the construct of malaise and somatoform psychopathology. External correlate data reported in Tellegen and Ben-Porath (2008/2011) provide support for the validity of MLS (Ben-Porath, 2012). Scores

on MLS were found to relate to a sense of poor health, sleep disruption, fatigue, and pain. Further, higher pre-surgical MLS scores in spine surgery candidates were one of several variables found to predict both higher levels of pain post-operatively and greater negative impacts of pain on the candidates' functioning (Marek, Block, & Ben-Porath, 2014).

Gastrointestinal Complaints, Head Pain Complaints, and Neurological Complaints. The next three scales were designed to assess somatic symptoms that could arise either as a function of a medical condition or in response to stress (Ben-Porath, 2012). Gastrointestinal Complaints (GIC) assesses somatic symptoms such as vomiting and poor appetite. Correlate data has revealed associations between GIC scores and reports of reduced appetite, gastric complaints, worry about health problems (Ben-Porath, 2012), and somatoform disorders (van der Heijden, Egger, Rossi, Grundel, & Derksen, 2013). Higher GIC scores have also been found to predict anxiety disorders (Haber & Baum, 2014). Head Pain Complaints (HPC) assesses head and neck pain; external correlates include headache complaints, additional somatic and pain symptoms, and concern about physical health (Ben-Porath, 2012). Finally, Neurological Complaints (NUC) assesses symptoms like dizziness and numbness. Correlate data revealed relationships with concentration difficulties, neurological complaints, sensory-motor dysfunction (Ben-Porath, 2012), and pain intensity in chronic low back pain patients (Tarescavage, Scheman, & Ben-Porath, 2014).

Cognitive Complaints. Cognitive Complaints (COG), the final scale in the Somatic/Cognitive set, relates to complaints about memory and concentration (Ben-Porath, 2012). COG was included in this set as a result of its high covariation with the

other Somatic/Cognitive Scales. COG scores were found to correlate with reports of memory complaints, concentration and cognition difficulties (Ben-Porath, 2012; Gervais, Ben-Porath, & Wygant, 2009, as cited in Ben-Porath, 2012), and task performance issues in police candidates (Tarescavage, Corey, & Ben-Porath, 2015); diagnoses of depression, somatoform disorders, and anxiety (Ben-Porath, 2012; van der Heijden et al., 2013); and even psychotic personality traits (Sellbom, Anderson, & Bagby, 2013).

The Internalizing Scales. Nine scales are included in this domain, which are divided into two sub-domains (Ben-Porath, 2012). The first four scales, which include Suicidal/Death Ideation (SUI), Helplessness/Hopelessness (HLP), Self-Doubt (SFD), and Inefficacy (NFC), are associated with the construct of demoralization (as assessed primarily by RCd). Furthermore, they are all also correlated with risk factors for suicide. The last five scales, which include Stress/Worry (STW), Anxiety (AXY), Anger Proneness (ANP), Behavior Restrictive Fears (BRF), and Multiple Specific Fears (MSF), are related to the construct of negative emotionality assessed by RC7 (Dysfunctional Negative Emotions).

Suicidal/Death Ideation. The first scale in this set, Suicidal/Death Ideation (SUI), was designed to assess suicidal ideation and past suicidal acts (Ben-Porath, 2012). This scale provides vital clinical data, as self-reports of suicidal ideation may occur even when such thoughts were not endorsed during in-person interviews (Glassmire, Stolberg, Greene, & Bongar, 2001, as cited in Ben-Porath, 2012). Correlate data revealed a host of associations between scores on SUI and relevant external criteria, including: past suicide attempts, persistent and current suicidal ideation, including the

presence of a suicide plan, hopelessness, and helplessness. Further, Gottfried, Bodell, Carbonell, and Joiner (2014) found support for the validity of the SUI scale, as evidenced by strong correlations with relevant external measures in a large (N = 998) outpatient sample.

Helplessness/Hopelessness. Helplessness/Hopelessness (HLP), the second scale, was designed to assess "pessimism about one's future prospects and the ability to improve them through self-change" (Ben-Porath, 2012, p. 112). Generally, both hopelessness and helplessness have been linked to suicide and self-injury. Additionally, hopelessness has been linked to depression and anxiety, poor responsiveness to antidepressants, Bipolar Disorder, and several medical conditions. External correlates for HLP scores include reports of helplessness and hopelessness, a strong relationship with Beck Hopelessness Scale (BHS; Beck & Steer, 1993) scores, suicidal ideation and attempts, diagnoses of depression (van der Heijden et al., 2013), and depressive personality traits (Sellbom et al., 2013).

Self-Doubt and Inefficacy. The next two scales, Self-Doubt (SFD) and Inefficacy (NFC), are discussed as a pair given their relatively high correlations with each other (Ben-Porath, 2012). Items in SFD relate to poor self-esteem and feeling inferior to others; NFC items relate to incapacitated decision-making skills in the face of emotional distress. The constructs of poor self-esteem and self-doubt relate to a number of negative mental health outcomes, including: depression and depressive personality traits (Ben-Porath, 2012; Haber & Baum, 2014; Sellbom et al., 2013), anxiety (van der Heijden et al., 2013), suicidal ideation (Ben-Porath, 2012), Posttraumatic Stress Disorder (PTSD), eating disorders (Ben-Porath, 2012; Tarescavage, Wygant, Boutacoff,

& Ben-Porath, 2013), and certain personality disorders and traits **(**Ben-Porath, 2012; Sellbom et al., 2013). SFD scores were found to correlate with feelings of insecurity and inferiority, self-doubt, and worthlessness; correlates for NFC included passivity, behavioral inhibition, and feelings of vulnerability.

Stress/Worry. Stress/Worry (STW), the first of the five Internalizing SP scales related to negative emotionality and the RC7 Scale, was designed to assess experiences of worry proneness, nervousness, and feeling pressured by time (Ben-Porath, 2012). STW items relate to demoralization, rumination, and excessive worry, constructs often related to depressive and anxiety disorders. STW scores were found to correlate with worry proneness and rumination (Ben-Porath, 2012; Brinker, Chin, & Wilkinson, 2014), nervousness (Ben-Porath, 2012), hopelessness, scores on measures assessing symptoms of anxiety and depression, diagnoses of depression and anxiety (van der Heijden et al., 2013), and Anxious, Avoidant, and Borderline Personality Disorder traits (Sellbom et al., 2013).

Anxiety. The Anxiety (AXY) scale was designed to assess persistent symptoms of anxiety characterized by frequent feelings of fright and experiences of nightmares (Ben-Porath, 2012). Furthermore, this scale taps the constructs of anxiety expectancy, the belief that certain stimuli will evoke anxiety, and anxiety sensitivity, the sense that the experience of anxiety will result in additional anxiety. Anxiety sensitivity in particular has been associated with PTSD. External correlate data revealed that scores on AXY were related to nightmares, intrusive ideation, fearfulness, anxiety disorder diagnoses, and symptoms of PTSD and depression.

Anger Proneness. Items on Anger Proneness (ANP) relate to irritability, difficulty controlling one's anger, and struggling with impatience in interpersonal situations (Ben-Porath, 2012). The construct assessed by ANP refers to anger as an affect state, distinguishing it from aggression and hostility. However, self-reports of anger have been found to predict assault in some populations, highlighting the importance of this scale. ANP scores were found to correlate with being angry and argumentative, having low frustration tolerance, hostility, past problems with juvenile misconduct (Ben-Porath, 2012), and Antisocial and Borderline Personality traits (Sellbom et al., 2013). Furthermore, Forbey et al. (2010) demonstrated good convergent validity between the ANP scale and the Anger Idioms Scale (Malgady, Rogler, & Cortes, 1996), a scale designed to assess how anger is manifested behaviorally.

Behavior Restrictive Fears and Multiple Specific Fears. The final two scales in the Internalizing set consist of Behavior Restrictive Fears (BRF) and Multiple Specific Fears (MSF; Ben-Porath, 2012). While BRF assesses fears "that inhibit and significantly restrict the individual's normal range of behaviors" (Ben-Porath, 2012, p. 114), MSF was designed to assess co-occurring phobias and specific kinds of fears. For example, MSF items assess fears related to animals and blood-injection-injury. Correlate data for BRF revealed associations with an assessment of Agoraphobia; data for MSF revealed positive correlations with harm avoidance and the amount of specific fears experienced. Further, more recent research by Phillips, Sellbom, Ben-Porath, and Patrick (2013) found a negative correlation between MSF scores and Psychopathic Personality Inventory (PPI; Lilienfeld & Andrews, 1996) scores in incarcerated samples of men and women.

The Externalizing Scales. Four scales comprise this domain, which are divided into two sub-domains (Ben-Porath, 2012). The scales associated with the first sub-domain, which include Juvenile Conduct Problems (JCP) and Substance Abuse (SUB), are conceptually related to Antisocial Behavior (RC4). The scales associated with the second sub-domain, which include Aggression (AGG) and Activation (ACT), are conceptually related to Hypomanic Activation (RC9). Elevations in each of these scales were found in a sample of probation violators as compared to a comparison sample of probation completers (Tarescavage, Luna-Jones, & Ben-Porath, 2014).

Juvenile Conduct Problems. Juvenile Conduct Problems (JCP), the first scale in this set, assesses for past patterns of juvenile misconduct, such as stealing (Ben-Porath, 2012). The construct assessed by JCP relates to juvenile Conduct Disorder, a diagnosis associated with increased chances of being treated on an inpatient basis, alcohol dependence, psychopathy, and interpersonal violence. External correlate data revealed associations between JCP scores and acting out behavior, stealing, truancy, difficulty with figures of authority, substance abuse, and with being diagnosed with Antisocial Personality Disorder. Furthermore, scores on JCP were found to be a strong predictor of Drug Court treatment non-completion (Mattson, Powers, Halfaker, Akeson, & Ben-Porath, 2012), premature termination from therapy (Anestis, Gottfried, & Joiner, 2015), and poor follow-up adherence to care for bariatric surgery patients (Tarescavage et al., 2013).

Substance Abuse. The second scale of the RC4 facets, Substance Abuse (SUB), was designed to assess use and abuse of alcohol and drugs (Ben-Porath, 2012). In contrast to substance use and abuse assessment on the MMPI-2, items on SUB are transparent; this

style has also been used successfully in a variety of other alcohol and drug assessment measures. As with JCP, assessing for substance abuse can prove important given its link to interpersonal violence. Correlations between SUB scores and relevant external criteria revealed strong associations with problems caused by substance abuse and substance abuse diagnoses (Ben-Porath, 2012; Haber & Baum, 2014; van der Heijden et al., 2013). Scores were also correlated with several external alcohol and drug abuse measures**.**

Aggression. Aggression (AGG), the first Externalizing SP Scale correlated with RC9, was created to assess violent behavior and physical aggression directed towards other individuals (Ben-Porath, 2012). The construct of aggression has been linked to interpersonal violence in inpatients and outpatients generally and in those diagnosed with Bipolar Disorder and PTSD specifically. Comparison between AGG scores and external measures revealed correlations with physical abusiveness, domestic violence, hostility, and homicidal ideation. Further, higher AGG scores are related to Antisocial Personality Disorder traits (Sellbom et al., 2013) and uncooperativeness (Tarescavage et al., 2015).

Activation. Activation (ACT), the final Externalizing SP Scale, was designed to assess symptoms typically found in Bipolar Disorder, including cycling moods, elation, overexcitation, and racing thoughts (Ben-Porath, 2012). Scores on other measures that also assess this construct have been linked to an increased chance of experiencing Bipolar Disorder, mania, hypomania, and substance abuse. External correlate data revealed correlations between ACT scores and hypomania, grandiose delusions, pressured speech, and a diagnosis of Bipolar Disorder. In addition to this correlate data

reported by Tellegen and Ben-Porath (2008/2011), ACT scores have been found to be significant predictors of Bipolar Disorder (Sellbom, Bagby, Kushner, Quilty, & Ayearst, 2012), the differential diagnosis between Major Depression and Bipolar Disorder (Watson, Quilty, & Bagby, 2010, as cited in Ben-Porath, 2012), impulsive personality traits (Sellbom et al., 2013), and, in a sample of police candidates, difficulty controlling one's behavior under duress (Tarescavage et al., 2015).

The Interpersonal Scales. The Interpersonal Scales are comprised of five individual scales designed to assess interpersonal functioning (Ben-Porath, 2012). They include Family Problems (FML), Interpersonal Passivity (IPP), Social Avoidance (SAV), Shyness (SHY), and Disaffiliativeness (DSF). SAV, SHY, and DSF are conceptually tied together around the construct of social isolation. As a whole, the ability of the Interpersonal Scales to accurately assess interpersonal problems has been strongly supported (Ayearst, Sellbom, Trobst, & Bagby, 2013).

Family Problems. Items answered in the keyed direction on Family Problems (FML) relate to family relationships marked by conflict and alienation (Ben-Porath, 2012). These relationships may relate to one's family of origin, current family, or both; FML does not allow the examiner to distinguish between these options. The construct of family dysfunction has been linked to a host of negative outcomes, including depression, substance abuse, physical health problems, personality disorder diagnoses, suicide attempts, and overeating (Ben-Porath, 2012; Tarescavage et al., 2013). Thus, it represents an important area of assessment. Correlations between FML scores and relevant external criteria, as reported by Tellegen and Ben-Porath (2008/2011) revealed

strong associations with hostility, alienation, familial discord, and therapist reports of families marked by the tendency to blame and resent each other.

Interpersonal Passivity. Interpersonal Passivity (IPP) was designed to assess an interpersonal style marked by unassertiveness and submissiveness (Ben-Porath, 2012). As with FML, interpersonal passivity has been linked to a variety of negative psychological outcomes, including decreased efficacy of social support and the fostering of dependency, a core construct of Dependent Personality Disorder. External correlations revealed strong associations between IPP scores and unassertive interpersonal styles, in addition to therapist reports of their patients being submissive and introverted.

Social Avoidance, Shyness, and Disaffiliativeness. As discussed above, Social Avoidance (SAV), Shyness (SHY), and Disaffiliativeness (DSF) are conceptually tied together around the construct of social isolation (Ben-Porath, 2012). Items answered in the keyed direction for SAV relate to a tendency for the individual to avoid social interaction, a core construct of Avoidant Personality Disorder. SAV scores were found to correlate positively with introversion and social fears. Forbey et al. (2010) also presented evidence supporting the convergent validity of SAV with the Social Avoidance Distress Scale (SADS; Watson & Friend, 1969), a measure of anxiety experienced in social situations. Negative correlations were reported for SAV scores with measures of warmth and social closeness (Ben-Porath, 2012)

Similarly, SHY was designed to assess "experiences of anxiety and discomfort associated with interacting with others" (Ben-Porath, 2012, p. 121). Research has found shyness to be highly heritable; it is also part of the diagnostic picture of Social Phobia.

Correlations between SHY scores and relevant extratest criteria revealed strong relationships with measures of anxiety, social fears, demoralization, and stress reactivity. Additionally, both SHY (van der Heijden et al., 2012) and SHY and DSF (Sellbom et al., 2013) scores were found to correlate strongly with a diagnosis of Avoidant Personality Disorder.

The DSF scale captures the tendency for an individual to want to be alone (Ben-Porath, 2012). This construct has been important in understanding the divide between Schizoid Personality Disorder, marked by a desire not to form relationships with others, and Avoidant Personality Disorder, which is characterized by a desire for interpersonal closeness and at the same time a debilitating sense of inferiority and fear of rejection. Correlates included depression, pessimism, social withdrawal, and suicidal ideation. Further, higher DSF scores have been found to correlate positively with assertiveness difficulties in police officer candidates (Tarescavage et al., 2015).

Summary. In reviewing these scales, two things become clear. First, the SP Scales are an incredibly rich source of clinical information. This is likely the result of their short and content-homogenous nature (Graham, 2012). Second, however, is these characteristics also suggest that these scales are vulnerable to invalid responding. In an examination of the effects of overreporting on the SP and PSY-5 Scales of the MMPI-2-RF, Burchett & Ben-Porath (2010) reported that instructing participants to intentionally feign psychopathology resulted in significant increases in SP Scale scores. Further, Dragon (2012) demonstrated the deleterious effects of increasing degrees of random responding on SP Scale interpretation. Thus, quantifying the effects of variable and fixed responding on the SP Scales would likely add to the existing literature in this area.

The Personality Psychopathology Five (PSY-5) Scales

Whereas the RC and SP Scales were designed to canvass Axis I pathology on the MMPI-2-RF, the PSY-5 Scales were intended to provide a dimensional model of Axis II pathology (Ben-Porath, 2012). The PSY-5 scales as they exist on the MMPI-2-RF were developed through four main steps. As with VRIN-r and TRIN-r, however, original forms of these scales existed prior to their inclusion on the MMPI-2 and MMPI-2-RF. The development of these scales is discussed followed by a description of their current form on the MMPI-2-RF.

PSY-5 model development. In the first step of development, Harkness (1992) set out to capture personality dimensions capable of describing both normal and abnormal personality. Harkness asked participants from non-clinical samples to group, based on similarity, items drawn from DSM-III-R (American Psychiatric Association [APA], 1987) criteria for personality disorders, markers of psychopathy identified by Cleckley (1982), and items drawn from Tellegen's MPQ. Harkness (1992) analyzed these groupings, resulting in a total of 60 clusters representing both disordered personality and normal personality functioning. Subsequently, Harkness and McNulty (1994) identified five latent factors from these clusters, labeling them the Personality Psychopathology Five (PSY-5).

Harkness and McNulty (1994) labeled the first of these dimensions Aggressiveness, which was designed to assess the tendency for individuals to employ instrumental aggression in interpersonal relationships (Ben-Porath, 2012; Harkness & McNulty, 2006). Second, Psychoticism related to the accuracy of an individual's reality testing skills. Third, Constraint (later relabeled and reverse keyed) assessed the degree of

control over one's behavior and harm avoidance. Fourth, Negative Emotionality/Neuroticism measured the tendency to experience anxiety, nervousness, and other negative emotions. Fifth, and finally, Positive Emotionality/Extroversion (also relabeled and reverse keyed) measured a disposition towards positive emotions and social experiences.

Transition to MMPI-2. In the second step of development, Harkness, McNulty, and Ben-Porath (1995) recognized the usefulness of the PSY-5 in assessing Axis II disorders (Ben-Porath, 2012). Using an analysis Harkness and colleagues referred to as replicated rational selection, MMPI-2 items were selected to represent the PSY-5 domains (Harkness & McNulty, 2006). The replicated rational selection procedure followed several steps. First, item selectors were trained in each of the PSY-5 domains. Following this training, selectors were asked to review the entire MMPI-2 item pool and select items reflective of that domain. Second, items were included in trial scales if over 51% of item selectors assigned an item to a PSY-5 domain and if the item did not relate to another domain. Third, and finally, the trial scales were assessed using large samples of college students and three clinical samples. Items were eliminated that evidenced poor correlations with the scale to which they were assigned and if they showed higher correlations with another trial scale.

Empirical Support. The resulting MMPI-2 PSY-5 Scales showed strong reliability as measured using the MMPI-2 normative sample (Harkness & McNulty, 2006). Furthermore, strong empirical support exists for the validity of these scales. First, the five-factor structure of the PSY-5 has been supported using several different techniques (Bagby, Ryder, Ben Dat, Bacchiochi, & Parker, 2002; Rouse, Finger, & Butcher, 1999)

and across both clinical and nonclinical samples (Rouse, 2007). Second, correlations between PSY-5 Scale scores and a host of related personality measures revealed strong support for convergent and discriminant validity (Ben-Porath, 2012; Harkness & McNulty, 2006). Third, associations between PSY-5 Scale scores and external criteria representing personality, internalizing, externalizing, and medical disorders provided strong support in the area of criterion validity (Ben-Porath, 2012; Harkness, Finn, McNulty, & Shields, 2012; Harkness & McNulty, 2006).

Transition to the MMPI-2-RF. Given the strong empirical support and usefulness of the MMPI-2 versions of the PSY-5 Scales, Harkness and McNulty were invited by Tellegen and Ben-Porath to update these scales for use with the MMPI-2-RF (Ben-Porath, 2012; Harkness et al., 2012). Of the original 136 items that comprised the PSY-5 Scales on the MMPI-2, Harkness and McNulty now had 96 items to use. Through internal psychometric and external criterion analyses, Harkness and McNulty (2007) removed and added items until arriving at a final item count of 104 items (Harkness et al., 2012). As with the MMPI-2 versions, the MMPI-2-RF versions of the PSY-5 scales were non-overlapping.

Empirical Support. Tellegen and Ben-Porath (2008/2011) reported strong reliability estimates for the revised PSY-5 Scales (Ben-Porath, 2012; Harkness et al., 2012), in addition to high correlations between the MMPI-2 and MMPI-2-RF versions of these scales. They also reported a series of external correlates for each scale using the same five validation samples referenced in the previous discussion of SP Scale correlates. Correlate information from this and other sources will be discussed below.

Scores on Aggressiveness-Revised (AGGR-r) correlated strongly with anger, interpersonal aggression, grandiosity, antisocial behavior (Ben-Porath, 2012), and uncooperativeness and poor functioning as a member of a team among police officer candidates (Tarescavage et al., 2015). Psychoticism-Revised (PSYC-r) scores were found to correlate with psychotic symptoms, magical ideation, perceptual aberration, and depression (Ben-Porath, 2012). Correlates for Disconstraint-Revised (DISC-r) were found to include substance abuse, narcissism, poor impulse control, and a history of juvenile misconduct. Among a sample of felons, higher scores on DISC-r were more characteristic of those who violated as compared to completed their parole (Tarescavage et al., 2014). Negative Emotionality/Neuroticism-Revised (NEGE-r) scores correlated strongly with depression, anxiety, hopelessness, and suicidal ideation (Ben-Porath, 2012). Finally, scores on Introversion/Low Positive Emotionality-Revised (INTR-r) were found to correlate with introversion, depression, hopelessness, and suicidal ideation. As a whole, Anderson et al. (2012) demonstrated strong patterns of convergence between each of the PSY-5 scales and the five respective domains of the 220-item Personality Inventory (PID-5; Krueger et al., 2012) for DSM-5 (APA, 2013).

Summary. As with the SP Scales, the MMPI-2-RF PSY-5 Scales provide a valuable source of information in the form of a dimensional assessment of personality. While these scales contain a greater number of items than the SP Scales, the PSY-5 Scales are nevertheless vulnerable to the deleterious effects of invalid responding. In the Burchett and Ben-Porath (2010) study discussed earlier, findings indicated that instructing participants to intentionally feign psychopathology resulted in significant scale score

increases. Thus, quantifying the effects of random and fixed responding on the PSY-5 Scales would likely add to the existing literature in this area.

CHAPTER III

RATIONALE AND HYPOTHESES

Rationale

From the literature reviewed above, it is clear that the MMPI-2-RF contains valuable sources of clinical information concerning psychopathology and personality. What is becoming clear through emerging research is that the interpretation of this information can be significantly distorted as a result of non-content-based responding. Therefore, the rationale for conducting this study is threefold.

First, Handel et al. (2010) and Dragon (2012) have collectively examined the effects of simulated responding on the interpretation MMPI-2-RF scales. While Handel and colleagues analyzed how RC Scale T-scores changed as a function of simulated random and fixed responding, Dragon examined the effects of varying degrees of random responding on the interpretation of H-O, RC, and SP Scales. These authors and others (e.g., Burchett & Ben-Porath, 2010) have called for an extension of these analyses into the SP and PSY-5 Scales. Therefore, this dissertation seeks to quantify the effects of random and fixed responding on mean T-scores of these scales, using the same data sets as Handel and colleagues. Results from this study will add to the existing literature base and clinical practice by providing interpretive recommendations reflective of varying degrees of random and fixed responding.

Second, Handel et al. (2010) and Dragon (2012) also examined the effects of simulated responding on convergent validity coefficients calculated between MMPI-2-RF scales and external measures. In the Handel and colleagues study, the Brief Psychiatric Rating Scale (BPRS; Overall & Gorham, 1988) was used as the external

measure. Therefore, this dissertation aims to extend these analyses to the PSY-5 and SP Scales. These analyses will illustrate the degradation of validity coefficients calculated between these scales and BPRS variables as a function of random and fixed responding.

Third, and as an extension of the second rationale, moderated multiple regression (MMR) analyses will be used to examine the extent to which VRIN-r (or TRIN-r) scores moderate the relationship between SP/PSY-5 Scale scores and BPRS scores. Burchett (2012), in an unpublished dissertation, used this procedure to examine how MMPI-2-RF overreporting validity indices (e.g., F-r) moderated the relationship between RC Scales and relevant extra-test measures. Burchett reported results indicating that moderation effects were present, and called for an extension of these analyses – including using VRIN-r and TRIN-r – to the SP and PSY-5 Scales. Thus, these analyses will be conducted to illustrate how random and fixed responding may moderate the relationship between SP/PSY-5 Scales and BPRS variables.

Hypotheses

The research by Handel et al. (2010) and Dragon (2012) are the only studies to date that have examined the effects of random and fixed responding on the substantive scales of the MMPI-2-RF. The results from Handel and colleagues, which focused exclusively on the RC Scales, found that while some scale mean T-scores increased as a result of random and fixed responding, this was not true for each scale. These findings were likely due to the differences in the relative number of items scored in either the true or false direction across scales. Similarly, the results from Dragon (2012) indicated that increasing degrees of simulated random responding resulted in an increase in most of the H-O, RC, and SP Scale mean T-scores across five experimental samples. However,

this study did not examine the effects of random responding on the PSY-5 Scales and fixed responding on the SP or PSY-5 Scales. Therefore, Hypothesis 1 states that increasing degrees of random and fixed responding will result in substantial score increases for a number of the SP and PSY-5 Scales. Further, it was proposed that differences in the effects of non-content-based invalid responding on mean scale T-scores would vary based on item keying for each scale (e.g., scales with most or all items keyed *true* would increase more rapidly with simulated acquiescent responding as compared to scales with most or all items keyed *false*).

Regarding the second rationale, both Handel et al. (2010) and Dragon (2012) demonstrated that increasing degrees of random and fixed responding for RC Scales, and increasing degrees of random responding for H-O, RC, and SP Scales, respectively, degraded convergent validity coefficients at simulated non-content-based responding rates greater than or equal to (≥) 30%. It should be noted that validity coefficients appeared relatively robust in both studies at rates below 30%. Therefore, Hypothesis 2a states that convergent validity coefficients for SP and PSY-5 Scales in the current study will degrade in a similar pattern under conditions of increasing simulated random and fixed responding. Specifically, these coefficients will: (a) be relatively robust to random insertion rates below 30% and (b) evidence substantial degradations at rates ≥ 30%. Finally, it should be noted that discriminant validity coefficient analyses will not be conducted due to the results of Handel and colleagues indicating that simulated non-content-based responding did not impact discriminant validity.

Regarding the third rationale, results from Burchett (2012) indicated that MMPI-2-RF overreporting validity indices moderated the relationship between RC Scales and

relevant extra-test measures. However, this author is not aware of any studies examining how non-content based validity indices (i.e., VRIN-r, TRIN-r) may serve as moderators between SP/PSY-5 Scales and extra-test measures. Therefore, very specific hypotheses concerning the moderating effects of these scales will not be proposed. However, in general, Hypothesis 2b states that increasing degrees of random and fixed responding will moderate the relationship between SP/PSY-5 Scales and BPRS variables.

CHAPTER IV

METHODS

Participants

This study uses the two data sets examined in Handel et al. (2010). The first of these data sets is the nongendered MMPI-2-RF normative sample (Ben-Porath & Forbey, 2003). The MMPI-2-RF normative sample (N = 2,276) was drawn from the normative sample created for the MMPI-2 (Ben-Porath, 2012; Handel et al., 2010). The primary difference is that the MMPI-2-RF sample contains an equal numberof men and women, resulting in a nongendered normative sample. Demographic information for the normative sample is provided in Table 1. Complete demographic information is available in Tellegen and Ben-Porath (2008/2011) and Ben-Porath & Forbey (2003).

The second data set is comprised of psychiatric inpatients (N = 704; Archer, Griffin, & Aiduk, 1995; Handel & Archer, 2008). Demographic information for the inpatient sample is also provided in Table 1. Participants with missing diagnoses (18.9%), followed by those with substance abuse or dependence diagnoses (17.3%), comprised the largest number of diagnoses as assessed by the DSM-III-R (APA, 1987). Additional diagnoses included major depressive disorder (16.6%), “other” diagnoses (15.8%), schizophrenia or psychotic disorders (12.9%), adjustment disorders (9.9%), and bipolar disorder (8.5%).

Table 1

Demographic Information for the Normative (N = 2,276) and Inpatient (N = 704) Samples

	Normative Sample		Inpatient Sample	
	#	(%)	#	(%)
Gender				
Male	1,138	(50.0)	363	(51.6)
Female	1,138	(50.0)	341	(48.4)
Race				
Caucasian	1,861	(81.8)	461	(65.4)
African American	264	(11.6)	144	(20.5)
Native American	71	(3.1)	0	(0.0)
Hispanic	67	(2.9)	0	(0.0)
Asian	13	(0.6)	0	(0.0)
Other	0	(0.0)	99	(14.1)

Note. # = the number of participants in each demographic variable; (%) = the percentage of participants from the respective total sample in each demographic variable; Other = either participant race that was not Caucasian or African American or participants who chose not to identify a race or ethnicity.

Measures

The MMPI-2-RF. The MMPI-2-RF consists of 338 items and a total of 51 scales (Ben-Porath, 2012). Of these 51 scales, 30 are examined in this dissertation. These include two validity scales (VRIN-r and TRIN-r), the 23 SP Scales, and the five PSY-5 Scales.

As discussed previously, reliability estimates for VRIN-r and TRIN-r in the MMPI-2-RF normative sample are relatively low (Tellegen & Ben-Porath, 2008/2011, as cited in Ben-Porath, 2012). Alpha coefficients for men and women on VRIN-r were .39 and .20, respectively; for TRIN-r, they were .37 and .23, respectively. Test-retest reliabilities, calculated from a subsample of the normative sample, were .52 and .40 for VRIN-r and TRIN-r, respectively. Finally, standard errors of measurement (SEM) of seven (VRIN-r) and eight (TRIN-r) T-score points were reported. Using the sample of mental health inpatients proposed for use in this study, Handel and Archer (2008) reported alpha coefficients for men and women on VRIN-r of .34 and .44, respectively; for TRIN-r, they reported estimates of .49 and .45, respectively.

With respect to the Somatic/Cognitive cluster of SP Scales in the normative sample, Tellegen and Ben-Porath (2008/2011, as cited in Ben-Porath, 2012) reported adequate reliability, noting that lower estimates were the result of the nonclinical normative sample producing scores indicative of range restriction. Alpha coefficients ranged from .52 (NUC) for men to .69 (GIC and COG) for women; test-retest reliabilities ranged from .54 to .82 for NUC and MLS, respectively. The Internalizing SP Scales evidenced similar psychometrics. Alpha coefficients ranged from .34 (SUI) for women to .72 (ANP) for men; test-retest estimates ranged from .65 to .85 for HLP and MSF,

respectively. The Externalizing Scales revealed relatively strong reliability estimates. Alpha coefficients ranged from .56 to .66 for JCP and AGG, respectively; test-rest estimates ranged from .77 to .87 for ACT and SUB, respectively. Lastly, alpha coefficients for the Interpersonal Scales ranged from .43 (DSF) for women to .78 (SAV) for men; test-retest estimates ranged from .60 to .88 for DSF and SHY, respectively.

SP and PSY-5 Scale alpha coefficients for male and female participants from the sample of mental health inpatients (Archer et al., 1995; Handel & Archer, 2008) were also calculated for this study. It should be noted that information regarding test-retest reliability was not available for this sample. Alpha coefficients ranged from .71 (NUC) for women to .85 (COG) for both genders for the Somatic/Cognitive SP Scales. For the Internalizing Scales, alpha coefficients ranged from .61 (STW) to .81 (SUI) for men; all of the Internalizing Scale alpha coefficients for women fell within this range. Alpha coefficients for the Externalizing Scales ranged from .68 (ACT) for women to .78 (SUB) for men. For the Interpersonal Scales, alpha coefficients ranged from .67 (DSF) to .83 (SAV) for women; all of the Externalizing Scale alpha coefficients for men fell within this range.

Finally, in the normative sample, PSY-5 scale alpha coefficients ranged from .69 (PSYC-r) for both genders to .78 (NEGE-r) for men; test-retest estimates ranged from .76 (PSYC-r) to .93 (DISC-r) (Tellegen & Ben-Porath, 2008/2011, as cited in Ben-Porath, 2012). SEM T-scores ranged from three to six. In the inpatient sample, alpha coefficients ranged from .71 (AGGR-r) to .92 (PSYC-r) for men; all of the PSY-5 Scale alpha coefficients for women fell within this range. As with the SP Scales, information regarding test-retest reliability was not available for this sample.

The BPRS. The BPRS (Overall & Gorham, 1988) is an 18-item scale designed to assess a wide variety of psychiatric symptoms, such as somatic concern, anxiety, and hostility (Faustman & Overall, 1999, as cited in Handel et al., 2010). The BPRS is clinician-rated, designed so that clinicians rank, on a scale of 1 (*not present*) to 7 (*extremely severe*), the degree to which each of the 18 symptoms exist in the patient. In one of the most comprehensive reviews of the studies available on the BPRS, Hedlund and Vieweg (1980, as cited in Handel et al., 2010) provided strong conclusions about the scale's reliability and validity. Regarding reliability, they reported interrater reliabilities of .67 to .88 for Blunted Affect and Hallucinatory Behavior, respectively. In support of validity, Hedlund and Vieweg found strong correlations between BPRS change scores and extratest data. For the purposes of the present study, the BPRS was included in the mental health inpatient sample as an external measure used to calculate criterion validity; it is used in the same manner for the present analyses.

CHAPTER V

PROCEDURES

The procedure for this study closely followed that used by Handel et al. (2010). Statistical analyses were performed using SPSS version 22. Because the analyses conducted in the current study are based on protocols gathered from the MMPI-2 normative sample (Butcher et al., 2001) and from a mental health inpatient sample (Archer et al., 1995), information concerning data collection can be found in these sources. It should be noted that protocols from the nongendered normative sample (Ben-Porath & Forbey, 2003) contain the 338 MMPI-2-RF items. However, the protocols from the inpatient data set contain the 567 items from the MMPI-2. Therefore, prior to beginning any of the inpatient analyses, the 229 MMPI-2 items that were not retained in the MMPI-2-RF item set were removed from each protocol. This resulted in the conversion of MMPI-2 protocols to MMPI-2-RF protocols, allowing for the same analyses to be conducted on both samples.

Regarding the application of validity criteria to protocols in both the normative and inpatient samples, two clusters were employed. For the analyses in which random responding, as measured by VRIN-r, would be simulated in increasing degrees, the following criteria were used: CNS less than (<) 15; TRIN-r < 80; F-r < 90; Fp-r < 80; L-r < 80; RBS < 80; FBS-r < 80; and Fs < 80 (Ben-Porath, 2012). For the analyses in which acquiescent and counter-acquiescent fixed responding, as measured by TRIN-r, would be simulated in increasing degrees, the following criteria were used: CNS < 15; VRIN-r < 80; F-r < 90; Fp-r < 80; L-r < 80; RBS < 80; FBS-r < 80; and Fs < 80. These criteria were used to eliminate protocols in which invalid responding was present or

"very likely." Further, chi square analyses were calculated for the inpatient sample to determine if rates of protocol invalidity differed based on participant race. It should be noted that chi square analyses were not conducted for the MMPI-2-RF normative sample, as the University of Minnesota Press did not provide information pertaining to participant race at the level of the individual case.

The experimental manipulation and data analyses proposed for this study followed closely the methods used by Handel et al. (2010). Using a computer simulation procedure, individual items from protocols in both samples were randomly replaced in increasing degrees to simulate random and both types of fixed responding (i.e., acquiescent and counter-acquiescent). The simulated responding ranged from 0% to 100% of items, increasing for iterations of analyses in increments of 10% (i.e., 20% simulated random responding, then 30%, 40%, etc.). Per Tellegen and Ben-Porath's (2008/2011) recommendation, unrounded, untruncated T-scores were used to calculate mean T-scores (Handel et al., 2010).

To quantify the effects of random and fixed responding on SP and PSY-5 Scales, and to examine Hypothesis 1, three indicators were used. First, scale mean T-scores that evidenced a change of ≥ 5 T-score points from baseline (0% random response insertion) were identified. This has previously been identified as a measure of clinically significant change (Ben-Porath, 2012). For each cluster of SP and PSY-5 Scales, a range of percentages, from lowest to highest, of random and fixed response insertion are presented to indicate the scale or scales that evidenced mean T-score changes of this magnitude. Second, 95% confidence intervals were calculated for SP and PSY-5 Scale mean T-scores across all levels (i.e., baseline to 90 and 100%) of random and fixed

responding. While non-overlapping confidence intervals clearly indicate statistically significant changes caused by increasing degrees of response insertion, it should be noted that the presence of a statistically significant difference is possible even if confidence intervals overlap with one other (Wolfe & Hanley, 2002). Third, and finally, SEM values were calculated in the baseline condition and served as the values to evaluate mean T-score deviations resulting from random and fixed response insertion. For each response insertion condition, multiples of SEM values that an experimental mean T-score (e.g., at 40% random response insertion) deviated from the baseline condition (i.e., 0% response insertion) were provided.

To quantify the effects of increasing degrees of random and fixed responding on the association between SP/PSY-5 Scales and relevant external criteria as represented by the BPRS, two sets of analyses were conducted. In examination of Hypothesis 2a, validity coefficients between scale scores and rationally selected BPRS items were calculated. These coefficients were calculated across all levels (i.e., 0% [baseline] to 100%) of random, fixed acquiescent, and fixed counter-acquiescent response insertion. To determine the magnitude of change in coefficients across levels of response insertion, Pearson *r*-values were squared and then compared with baseline values (Handel et al., 2010). Given that not all SP and PSY-5 Scales are represented by BPRS domains, only those scales for which there was a representative correlate were included in these calculations. Therefore, the following proposed pairs are presented with the MMPI-2-RF first, followed the by BPRS item. Additional scale pairs were added based on an inspection of resulting validity coefficients. The proposed pairs included:

1. Malaise (MLS) - Somatic Concern

2. Anxiety (AXY) - Anxiety
3. Psychoticism-Revised (PSYC-r) - Conceptual Disorganization
4. Activation (ACT) - Grandiosity
5. Helplessness/Hopelessness (HLP) - Depressive Mood
6. Aggression (AGG) - Hostility
7. PSYC-r - Hallucinatory Behavior
8. PSYC-r - Unusual Thought Content

In examination of Hypothesis 2b, several steps were performed. First, inpatient data sets (with validity criteria applied) were prepared to represent random, fixed acquiescent, and fixed counter-acquiescent responding conditions. Using the random responding condition as an example, the representative data was created by: (a) randomly selecting 50% of cases, (b) inserting random responses into 80% of items of one half of the data set using the computer simulation procedure described previously, and (c) recombining the two halves of the data set in preparation for the MMR analyses (described below). This simulated non-content-based responding procedure was used because the percentages of elevated VRIN-r and TRIN-r scores were relatively low in the inpatient sample. This finding was expected given that protocols were collected from participants who were administered the MMPI-2 under standard instructions. Second, regression diagnostics were evaluated.

Third, a series of four steps were used for conducting the MMR analyses. This procedure follows closely that used by Burchett (2012) and originally developed by Aguinis, Culpepper, and Pierce (2010; cited in Burchett, 2012). First, interaction terms

were created for each MMR by multiplying the raw scores of the proposed predictor variable (e.g., MLS) with the raw scores of the proposed moderator variable (e.g., VRIN-r). Second, a linear regression was performed with the predictor variable entered in the first step, the moderator variable and the interaction term entered in the second step, and the criterion measure (e.g., BPRS1: Somatic Concern) entered as the dependent variable. This linear regression was performed to determine if there was a significant moderating effect among these variables. The final two steps were performed if the results of the initial linear regression were significant. In the third step, a linear regression was performed with the predictor and moderator variable entered in the first step, the interaction term entered in the second step, and the criterion measure entered as the dependent variable. This step was conducted to determine if slope differences were present. In the fourth and final step, a linear regression was performed with the predictor variable in the first step and the moderator variable entered in the second step. This was conducted to examine the presence of intercept differences. Statistical significance values and changes in R^2 effect sizes were reported. It should be noted that prior to creating the interaction term for each MMR, raw scores were not centered. Kromrey and Foster-Johnson (1998) reported results indicating that centered versus uncentered raw scores are equivalent in terms of regression coefficients and model significance as assessed by R^2.

Finally, MMR power analysis tables provided by Aiken and West (1991) were reviewed to determine the appropriate sample size. Given that there have been no prior studies examining the moderating effects of non-content-based responding on MMPI-2-RF criterion validity, a medium effect size was estimated. Thus, at 80% power, a

medium effect size, and .70 predictor reliabilities, a sample size of 192 was required (Aiken & West, 1991).

CHAPTER VI

RESULTS

The results provided below are grouped into two primary clusters. First, and in examination of Hypothesis 1, the effects of simulated random and fixed responding on SP and PSY-5 Scale mean T-scores from the normative and inpatient samples are presented. Second, and in examination of Hypotheses 2a and 2b, external validity analyses from the inpatient sample are reviewed.

MMPI-2-RF Normative Sample Analyses

Validity criteria were applied prior to beginning the random, fixed acquiescent, and fixed counter-acquiescent insertion analyses for the MMPI-2-RF normative sample. For the random insertion analyses, the following criteria used were: CNS < 15; TRIN-r < 80; F-r < 90; Fp-r < 80; L-r < 80; RBS < 80; FBS-r < 80; and Fs < 80 (Ben-Porath, 2012). For the fixed insertion analyses, the following criteria were used: CNS < 15; VRIN-r < 80; F-r < 90; Fp-r < 80; L-r < 80; RBS < 80; FBS-r < 80; and Fs < 80. These conservative criteria were applied in order to examine the effects of simulated random and fixed responding in the absence of the possible confounding effects of even moderate levels of other invalidating response styles (i.e., overreporting or underreporting). Table 2 displays the number of protocols eliminated from the total normative sample (N = 2,276) by a sequential application of individual validity criteria. Table 3 illustrates the frequency, as represented by a numerical count and percentage, of protocols from the total normative sample that would be identified as invalid by individual validity criteria. It should be noted that the number of protocols identified in Table 3 exceeded the number of protocols excluded by the sequential application of

validity criteria represented in Table 2. This is due to the fact that eliminating protocols sequentially (e.g., VRIN-r, then F-r) reduces the number of protocols that can be eliminated subsequently. The application of these criteria resulted in sample sizes of n = 2,110 and n = 2,124 for VRIN-r and TRIN-r, respectively. Frequency counts of each validity measure were calculated after the application of these criteria to confirm that there were not any invalid protocols remaining.

Table 2

Protocol Elimination From the Normative Sample (N = 2,276) as a Function of Sequential Application of MMPI-2-RF Validity Criteria

Validity Scale Name	Random Insertion Condition #	Fixed Insertion Condition #
CNS	5	5
VRIN-r	--	20
TRIN-r	38	--
F-r	14	14
Fp-r	27	27
Fs	26	29
FBS-r	17	16
RBS	9	11
L-r	30	30
Total	166	152
Resulting Sample Size	2,110	2,124

Note. # = number of protocols eliminated by each validity criteria; CNS = Cannot Say; VRIN-r = Variable Response Inconsistency – Revised; TRIN-r = True

Table 2 Continued

Response Inconsistency – Revised; F-r = Infrequent Responses – Revised; Fp-r = Infrequent Psychopathology Responses – Revised; Fs = Infrequent Somatic Responses; FBS-r = Symptom Validity – Revised; RBS = Response Bias; L-r = Uncommon Virtues – Revised. [--] = The validity criterion was not applied to this condition.

Table 3

The Frequency of Invalid Protocols in the Normative Sample (N = 2,276) as Identified by MMPI-2-RF Validity Criteria

Validity Scale Name	Protocols Identified #	(%)
CNS	5	(0.002)
VRIN-r	20	(0.009)
TRIN-r	39	(0.017)
F-r	14	(0.006)
Fp-r	35	(0.015)
Fs	39	(0.017)
FBS-r	26	(0.011)
RBS	24	(0.011)
L-r	34	(0.015)

Note. # = number of protocols eliminated by each validity criteria; (%) = percentage of the total number of protocols identified as invalid by each validity criteria; CNS = Cannot Say; VRIN-r = Variable Response Inconsistency – Revised; TRIN-r = True Response Inconsistency – Revised; F-r = Infrequent Responses – Revised; Fp-r =

Table 3 Continued

Infrequent Psychopathology Responses – Revised; Fs = Infrequent Somatic Responses; FBS-r = Symptom Validity – Revised; RBS = Response Bias; L-r = Uncommon Virtues – Revised.

Random response insertion. The results of increasing degrees of simulated random responding on SP and PSY-5 Scale mean T-scores are presented in Tables 1 through 6. It should be noted that these tables include SP and PSY-5 Scale mean T-scores for baseline (i.e., 0% insertion for random, acquiescent, and counter-acquiescent responding) and response insertion conditions (i.e., 10-100% insertion for random, acquiescent, and counter-acquiescent responding). Only the results for the random response insertion analyses will be discussed in this section; results of acquiescent and counter-acquiescent responding will be discussed subsequently. These tables also include mean T-score standard deviations; alphabetical superscripts to indicate the magnitude of the deviation, as indicated by multiples of SEMs, between a SP or PSY-5 Scale baseline mean T-score and a response insertion mean T-score (e.g., the mean T-score for the SP Scale MLS at 50% random response insertion deviated by two SEMs from baseline); and 95% confidence intervals for each mean T-score. Further, should be noted that the Internalizing SP scales are divided into two separate tables: Table 2 presents the SP Scales associated with the construct of demoralization, as represented by the RCd Scale, while Table 3 presents the SP Scales associated with the construct of dysfunctional negative emotions, as represented by the RC7 Scale.

For the 28 SP and PSY-5 Scales, increasing degrees of simulated random responding resulted in a monotonic increase in scale mean T-scores. However, these SP and PSY-5 mean T-scores differed in their susceptibility to score distortion. These differences are discussed below, using the three indicators described in the Procedures section.

Table 4

The Effects of Increasing Degrees of Variable, Fixed Acquiescent, and Fixed Counter-Acquiescent Response Insertion on Normative Sample Mean Somatic/Cognitive Scale T-Scores, Standard Deviations, and 95% Confidence Intervals.

Response insertion	MLS			GIC			HPC			NUC			COG		
percentage	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
0% (R)	49.7	9.3	[49.3, 50.1]	49.3	9.1	[48.9, 49.7]	49.4	9.5	[49.0, 49.8]	49.3	9.4	[48.9, 49.7]	49.3	9.3	[48.9, 49.7]
0% (T or F)	49.8	9.3	[49.4, 50.2]	49.3	9.1	[49.0, 49.7]	49.4	9.5	[49.0, 49.8]	49.3	9.4	[48.9, 49.7]	49.3	9.4	[48.9, 49.7]
10% (R)	51.2	9.0	[50.8, 51.6]	53.0	10.8	[52.6, 53.5]	51.3	9.3	[50.9, 51.7]	52.9	9.6	[52.5, 53.3]	51.7	9.0	[51.3, 52.1]
10% (T)	49.4	8.9	[49.0, 49.8]	55.6	11.5[a]	[55.1, 56.1]	51.2	9.5	[50.8, 51.6]	51.3	9.3	[50.9, 51.7]	53.8	8.9	[53.4, 54.1]
10% (F)	53.5	9.1	[53.1, 53.9]	51.0	9.8	[50.5, 51.3]	51.7	9.2	[51.3, 52.1]	55.0	9.4	[54.6, 55.4]	49.9	8.9	[49.5, 50.3]
20% (R)	52.7	8.9	[52.3, 53.1]	56.5	11.6[a]	[56.1, 57.0]	53.3	9.1	[52.9, 53.7]	55.9	9.4	[55.5, 56.3]	54.0	8.4	[53.6, 54.4]
20% (T)	48.9	8.5	[48.5, 49.3]	60.3	11.8[a]	[59.8, 60.8]	53.1	9.2	[52.7, 53.5]	53.1	9.0	[52.7, 53.5]	57.2	8.7[a]	[56.8, 57.5]
20% (F)	56.1	8.8[a]	[55.7, 56.4]	52.0	10.1	[51.5, 52.4]	53.1	8.9	[52.7, 53.5]	58.7	8.9[a]	[58.4, 59.1]	50.4	8.2	[50.0, 50.7]
30% (R)	53.7	8.5	[53.3, 54.0]	59.8	11.8[a]	[59.3, 60.3]	54.9	9.0	[54.5, 55.2]	58.9	9.0[a]	[58.6, 59.3]	56.0	8.3[a]	[55.7, 56.4]
30% (T)	48.6	7.8	[48.2, 48.9]	64.9	11.4[b]	[64.4, 65.4]	55.1	8.6[a]	[54.7, 55.5]	54.8	8.7	[54.4, 55.2]	60.7	8.4[b]	[60.4, 61.1]

Table 4 Continued

Response insertion percentage	MLS			GIC			HPC			NUC			COG		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
30% (F)	59.1	8.6[a]	[58.8, 59.5]	53.2	10.3	[52.8, 53.7]	55.3	8.2[a]	[55.0, 55.7]	62.7	8.6[a]	[62.3, 63.0]	50.9	7.8	[50.5, 51.2]
40% (R)	55.2	8.2	[54.8, 55.5]	62.3	11.8[b]	[61.8, 62.8]	56.8	8.6[a]	[56.4, 57.2]	61.7	8.8[a]	[61.3, 62.0]	58.1	8.1[a]	[57.8, 58.5]
40% (T)	48.3	7.5	[48.0, 48.6]	69.1	10.6[c]	[68.7, 69.6]	56.7	8.2[a]	[56.3, 57.0]	56.4	8.3[a]	[56.1, 56.8]	64.0	8.4[b]	[63.6, 64.3]
40% (F)	62.2	8.0[b]	[61.8, 62.5]	55.0	10.2[a]	[54.6, 55.4]	56.5	7.5[a]	[56.2, 56.8]	66.1	7.9[b]	[65.8, 66.5]	51.4	7.4	[51.1, 51.7]
50% (R)	56.3	8.1[a]	[56.0, 56.7]	65.0	11.4[b]	[64.5, 65.5]	58.3	8.6[a]	[57.9, 58.6]	64.1	8.5[b]	[63.7, 64.4]	59.5	8.4[a]	[59.1, 59.9]
50% (T)	47.8	7.0	[47.5, 48.1]	72.7	9.5[d]	[72.3, 73.1]	58.3	7.5[a]	[58.0, 58.7]	58.0	7.3[a]	[57.7, 58.3]	67.6	8.0[c]	[67.3, 67.9]
50% (F)	65.3	7.7[b]	[65.0, 65.6]	56.4	9.9[a]	[56.0, 56.8]	58.4	6.7[a]	[58.1, 58.7]	69.5	7.5[c]	[69.2, 69.8]	52.0	6.5	[51.7, 52.3]
60% (R)	58.1	8.0[a]	[57.8, 58.5]	66.8	11.5[c]	[66.3, 67.3]	59.4	8.4[a]	[59.1, 59.8]	66.2	8.5[b]	[65.8, 66.6]	61.8	8.2[b]	[61.4, 62.1]
60% (T)	47.7	6.2	[47.4, 47.9]	75.6	8.9[d]	[75.2, 75.9]	59.5	7.0[a]	[59.2, 59.8]	59.3	6.6[a]	[59.1, 59.6]	71.1	7.9[c]	[70.8, 71.4]
60% (F)	67.9	7.2[c]	[67.6, 68.2]	57.9	9.4[a]	[57.5, 58.3]	59.8	6.2[a]	[59.5, 60.0]	72.9	7.1[c]	[72.6, 73.2]	52.5	5.8	[52.2, 52.7]
70% (R)	59.4	7.9[a]	[59.0, 59.7]	69.7	10.9[c]	[69.2, 70.1]	61.2	8.3[b]	[60.9, 61.6]	68.4	8.4[b]	[68.1, 68.8]	63.4	8.1[b]	[63.0, 63.7]
70% (T)	47.1	5.6	[46.9, 47.4]	79.0	7.5[e]	[78.7, 79.3]	61.0	6.0[b]	[60.7, 61.3]	61.0	5.9[a]	[60.8, 61.3]	74.7	7.0[d]	[74.4, 75.0]
70% (F)	70.8	6.5[c]	[70.5, 71.0]	59.3	8.6[a]	[58.9, 59.7]	61.1	5.3[a]	[60.9, 61.3]	76.1	6.3[c]	[75.8, 76.4]	53.1	4.9	[52.9, 53.3]

Table 4 Continued

Response insertion percentage	MLS *M*	MLS *SD*	MLS *95% CI*	GIC *M*	GIC *SD*	GIC *95% CI*	HPC *M*	HPC *SD*	HPC *95% CI*	NUC *M*	NUC *SD*	NUC *95% CI*	COG *M*	COG *SD*	COG *95% CI*
80% (R)	60.7	8.1[a]	[60.4, 61.1]	71.8	10.5[d]	[71.4, 72.3]	62.2	8.3[b]	[61.8, 62.6]	70.9	8.3[c]	[70.5, 71.2]	65.5	8.5[b]	[65.1, 65.9]
80% (T)	47.0	4.8	[46.8, 47.2]	82.0	6.5[e]	[82.1, 82.6]	62.4	4.9[b]	[62.1, 62.6]	62.1	4.6[a]	[61.9, 62.3]	78.4	5.9[e]	[78.1, 78.6]
80% (F)	74.1	5.5[d]	[73.8, 74.3]	60.8	7.4[b]	[60.5, 61.1]	62.4	4.4[b]	[62.2, 62.6]	79.1	5.4[d]	[78.9, 79.4]	53.5	3.7	[53.3, 53.6]
90% (R)	61.7	7.8[b]	[61.4, 62.1]	73.9	10.3[d]	[73.4, 74.3]	64.0	8.1[b]	[63.6, 64.3]	72.5	8.4[c]	[72.2, 72.9]	67.6	8.3[c]	[67.2, 67.9]
90% (T)	47.8	3.3	[46.6, 46.9]	85.3	4.8[f]	[85.1, 85.5]	63.8	3.6[b]	[63.7, 64.0]	63.5	3.3[b]	[63.3, 63.6]	82.0	4.2[e]	[81.8, 82.2]
90% (F)	77.0	4.0[d]	[76.8, 77.1]	62.3	5.4[b]	[62.1, 62.5]	63.9	3.0[b]	[63.8, 64.1]	82.5	3.8[d]	[82.3, 82.7]	54.0	2.7	[53.9, 54.1]
100% (R)	63.1	7.5[b]	[62.8, 63.4]	76.3	9.4[d]	[75.9, 76.7]	65.1	8.1[b]	[64.8, 65.5]	75.1	8.4[c]	[74.7, 75.5]	69.6	8.3[c]	[69.3, 70.0]
100% (T)	46.5	0.0	[--, --]	88.3	0.0[f]	[--, --]	65.1	0.0[b]	[--, --]	64.6	0.0	[--, --]	85.6	0.0[f]	[--, --]
100% (F)	79.0	0.0[e]	[--, --]	63.8	0.0[b]	[--, --]	65.1	0.0[b]	[--, --]	85.6	0.0[e]	[--, --]	54.4	0.0	[--, --]

Note. n = 2,110 and n = 2,124 for variable and acquiescent/counter-acquiescent response insertion, respectively. MLS = Malaise;

Table 4 Continued

GIC = Gastrointestinal Complaints; HPC = Head Pain Complaints; NUC = Neurological Complaints; COG = Cognitive Complaints; R = random; T = true; and F = false.

[a,b,c,d,e,f] The mean score differs from the 0% variable, acquiescent, or counter-acquiescent response insertion baseline mean score by one, two, three, four, five, and six times this scale's standard error of measurement, respectively.

When fixed acquiescent and counter-acquiescent responding reaches 100%, T-scores become constant and the standard deviations equal zero.

The format of 95% confidence interval results is as follows: Mean [lower, upper bounds of 95% Confidence Interval]. [--, --] = The confidence interval for this mean score could not be calculated due to the same reason described above. The non-gendered normative sample, reported in Tellegen & Ben-Porath (2008), is the basis for all of the results presented in this table.

Table 5

The Effects of Increasing Degrees of Variable, Fixed Acquiescent, and Fixed Counter-Acquiescent Response Insertion on Normative Sample Mean Internalizing (RCd-Associated) Scale T-Scores, Standard Deviations, and 95% Confidence Intervals.

Response insertion	SUI			HLP			SFD			NFC		
percentage	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
0% (R)	49.4	9.1	[49.0, 49.7]	49.6	9.7	[49.1, 50.0]	49.5	9.6	[49.1, 49.9]	49.6	9.7	[49.2, 50.0]
0% (T or F)	49.4	9.1	[49.0, 49.8]	49.5	9.6	[49.1, 50.0]	49.5	9.6	[49.1, 49.9]	49.6	9.8	[49.2, 50.0]
10% (R)	53.6	11.7	[53.1, 54.1]	51.2	9.8	[50.8, 51.6]	50.6	9.2	[50.2, 50.9]	50.3	9.1	[49.9, 50.7]
10% (T)	58.3	13.2[a]	[57.7, 58.8]	52.7	10.0	[52.2, 53.1]	51.9	9.5	[51.5, 52.3]	52.4	9.3	[52.0, 52.8]
10% (F)	49.1	8.7	[48.7, 49.4]	49.7	9.5	[49.3, 50.1]	48.6	8.8	[48.2, 49.0]	48.2	8.9	[47.9, 48.6]
20% (R)	57.0	13.1	[56.5, 57.6]	52.8	9.7	[52.4, 53.2]	51.3	9.0	[50.9, 51.7]	51.2	8.4	[50.9, 51.6]
20% (T)	66.3	14.8[b]	[65.6, 66.9]	55.5	10.3	[55.1, 55.9]	54.3	9.3	[54.0, 54.8]	55.0	8.8[a]	[54.6, 55.3]
20% (F)	48.6	8.1	[48.2, 48.9]	49.9	9.2	[49.5, 50.3]	48.1	9.2	[47.7, 48.4]	47.1	8.4	[46.8, 47.5]
30% (R)	61.8	14.1[a]	[61.2, 62.4]	54.4	10.1	[53.9, 54.8]	52.2	8.5	[51.9, 52.6]	52.0	7.8	[51.6, 52.3]

Table 5 Continued

Response insertion percentage	SUI			HLP			SFD			NFC		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
30% (T)	73.2	15.1[c]	[72.5, 73.8]	58.5	10.2[a]	[58.1, 59.0]	56.6	9.2[a]	[56.2, 57.0]	57.8	8.6[a]	[57.4, 58.1]
30% (F)	48.2	7.6	[47.9, 48.5]	50.1	8.8	[49.7, 50.4]	47.3	7.5	[47.0, 47.6]	46.0	7.5	[45.6, 46.3]
40% (R)	65.1	14.5[b]	[64.5, 65.7]	55.7	10.0	[55.3, 56.2]	52.9	8.2	[52.5, 53.2]	52.6	7.2	[52.3, 52.9]
40% (T)	79.9	15.0[c]	[79.3, 80.5]	61.6	10.2[a]	[61.2, 62.1]	59.1	9.1[a]	[58.7, 59.5]	60.6	8.4[b]	[60.3, 61.0]
40% (F)	47.9	7.2	[47.5, 48.2]	50.1	8.3	[49.7, 50.4]	46.4	6.8	[46.2, 46.7]	44.6	7.0	[44.3, 44.9]
50% (R)	68.5	14.9[b]	[67.8, 69.1]	56.9	10.2	[56.5, 57.4]	53.5	8.2	[53.2, 53.9]	53.2	7.1	[52.9, 53.5]
50% (T)	86.5	14.5[d]	[85.8, 87.1]	65.2	9.5[b]	[63.8, 64.6]	61.4	9.1[b]	[61.0, 61.8]	63.1	8.3[b]	[62.7, 63.4]
50% (F)	47.4	6.4	[47.1, 47.7]	50.5	7.8	[50.1, 50.8]	45.8	6.1	[45.6, 46.1]	43.4	6.4	[43.1, 43.7]
60% (R)	72.7	15.5[c]	[72.0, 73.3]	58.6	10.2[a]	[58.2, 59.1]	54.4	7.8	[54.1, 54.7]	53.8	7.6	[53.5, 54.1]
60% (T)	92.0	14.0[e]	[91.4, 92.6]	67.1	9.0[b]	[66.7, 67.5]	63.6	8.6[b]	[63.3, 64.0]	66.7	7.9[c]	[66.4, 67.0]
60% (F)	47.1	6.0	[46.8, 47.3]	50.5	7.1	[50.2, 50.8]	44.9	5.3	[44.7, 45.1]	42.4	5.9[a]	[42.2, 42.7]
70% (R)	75.7	15.3[c]	[75.0, 76.3]	60.4	10.1[a]	[60.0, 60.9]	55.1	7.7[a]	[54.7, 55.4]	54.7	6.4	[54.4, 54.9]
70% (T)	97.8	13.0[f]	[97.2, 98.3]	70.0	8.3[b]	[69.7, 70.4]	66.6	8.1[c]	[66.3, 67.0]	70.2	7.1[c]	[69.9, 70.5]

Table 5 Continued

Response insertion percentage	SUI			HLP			SFD			NFC		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
70% (F)	46.6	5.0	[46.4, 46.8]	50.7	6.3	[50.4, 51.0]	44.4	4.9	[44.2, 44.6]	40.8	5.2[a]	[40.6, 41.0]
80% (R)	78.6	14.5[c]	[77.9, 79.2]	61.6	10.4[a]	[61.2, 62.1]	56.0	7.4[a]	[55.6, 56.3]	55.5	6.5[a]	[55.2, 55.7]
80% (T)	104.3	11.3[g]	[103.9, 104.8]	72.8	7.1[c]	[72.5, 73.1]	69.7	7.2[c]	[69.4, 70.0]	73.5	5.9[d]	[73.3, 73.8]
80% (F)	46.1	4.0	[45.9, 46.3]	50.8	5.3	[50.6, 51.0]	43.6	4.1[a]	[43.4, 43.8]	39.4	4.5[a]	[39.2, 39.6]
90% (R)	81.7	14.9[d]	[81.1, 82.4]	63.1	10.5[a]	[62.7, 63.6]	56.8	7.7[a]	[56.4, 57.1]	56.3	6.4[a]	[56.0, 56.6]
90% (T)	110.5	8.2[g]	[110.1, 110.8]	75.7	5.1[c]	[75.5, 75.9]	72.4	5.5[d]	[72.2, 72.7]	76.9	4.3[e]	[76.7, 77.1]
90% (F)	45.7	2.8	[45.6, 45.8]	51.3	3.6	[51.1, 51.4]	42.7	2.9[a]	[42.6, 42.8]	37.9	3.4[a]	[37.7, 38.0]
100% (R)	84.7	14.5[d]	[84.1, 85.0]	64.6	10.4[b]	[64.1, 65.0]	57.5	7.8[a]	[57.2, 57.9]	57.0	6.6[a]	[56.8, 57.3]
100% (T)	116.5	0.0[h]	[--, --]	78.5	0.0[c]	[--, --]	75.7	0.0[d]	[--, --]	80.4	0.0[e]	[--, --]
100% (F)	45.4	0.0	[--, --]	51.6	0.0	[--, --]	41.8	0.0[a]	[--, --]	36.1	0.0[b]	[--, --]

Note. n = 2,110 and n = 2,124 for variable and acquiescent/counter-acquiescent response insertion,

Table 5 Continued

respectively. RCd =

Demoralization; SUI = Suicidal/Death Ideation; HLP =Helplessness/Hopelessness; SFD = Self-Doubt;

NFC = Inefficacy; R = random; T = true; and F = false.

[a,b,c,d,e,f,g,h] The mean score differs from the 0% variable, acquiescent, or counter-acquiescent baseline mean score by one, two, three, four, five, six, seven, and eight times this scale's standard error of measurement, respectively.

When fixed acquiescent and counter-acquiescent responding reaches 100%, T-scores become constant and the standard deviations equal zero.

The format of 95% confidence interval results is as follows: Mean [lower, upper bounds of 95% Confidence Interval]. [--, --] = The confidence interval for this mean score could not be calculated due to the same reason described above. The non-gendered normative sample, reported in Tellegen & Ben-Porath (2008), is the basis for all of the results presented in this table.

Table 6

The Effects of Increasing Degrees of Variable, Fixed Acquiescent, and Fixed Counter-Acquiescent Response Insertion on Normative Sample Mean Internalizing (RC7-Associated) Scale T-Scores, Standard Deviations, and 95% Confidence Intervals.

Response insertion	STW			AXY			ANP			BRF			MSF		
percentage	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
0% (R)	49.6	9.7	[49.2, 50.0]	49.3	9.2	[48.9, 49.7]	49.6	9.8	[49.2, 50.1]	49.4	9.5	[49.0, 49.8]	49.7	9.8	[49.3, 50.1]
0% (T or F)	49.5	9.7	[49.1, 50.0]	49.3	9.2	[48.9, 49.7]	49.6	9.8	[49.2, 50.1]	49.4	9.4	[49.0, 49.8]	49.7	9.9	[49.3, 50.2]
10% (R)	50.2	9.3	[49.8, 50.6]	52.3	10.4	[51.9, 52.7]	50.6	9.1	[50.2, 51.0]	53.8	10.4	[53.3, 54.2]	50.1	9.1	[49.7, 50.5]
10% (T)	50.9	9.4	[50.5, 51.3]	55.7	11.4	[55.2, 56.2]	51.5	9.2	[51.1, 51.9]	57.0	10.7[a]	[56.5, 57.4]	50.0	9.0	[49.6, 50.4]
10% (F)	49.3	9.0	[48.9, 49.7]	48.7	8.6	[48.3, 49.0]	49.8	9.1	[49.4, 50.1]	49.9	9.4	[49.5, 50.3]	50.5	9.0	[50.1, 50.9]
20% (R)	50.7	8.8	[50.3, 51.0]	55.2	11.1	[54.8, 55.7]	51.4	8.6	[51.1, 51.8]	57.5	10.7[a]	[57.0, 57.9]	50.4	8.3	[50.1, 50.8]
20% (T)	52.4	9.2	[52.0, 52.8]	61.7	12.4[a]	[61.2, 62.2]	53.0	8.6	[52.7, 53.4]	63.5	11.2[b]	[63.0, 64.0]	50.2	8.3	[49.9, 50.6]
20% (F)	49.2	8.6	[48.8, 49.5]	48.3	8.2	[47.9, 48.6]	49.8	8.1	[49.4, 50.1]	50.6	9.3	[50.3, 51.0]	50.9	8.4	[50.5, 51.2]
30% (R)	51.3	8.7	[51.0, 51.7]	57.3	11.5[a]	[56.9, 57.8]	52.3	8.2	[52.0, 52.7]	60.5	10.7[a]	[60.0, 60.9]	50.8	7.9	[50.5, 51.2]
30% (T)	53.7	8.9	[53.3, 54.1]	66.6	12.5[b]	[66.1, 67.2]	54.4	8.1	[54.1, 54.8]	69.7	11.0[b]	[69.2, 70.2]	50.3	7.6	[50.0, 50.7]

Table 6 Continued

Response insertion percentage	STW			AXY			ANP			BRF			MSF		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
30% (F)	48.9	7.9	[48.6, 49.3]	47.9	7.6	[47.5, 48.2]	49.9	7.7	[49.5, 50.2]	51.1	8.9	[50.7, 51.5]	51.4	7.5	[51.0, 51.7]
40% (R)	51.9	8.3	[51.5, 52.2]	60.2	11.9[a]	[59.7, 60.7]	52.9	7.6	[52.6, 53.2]	64.2	11.2[b]	[63.7, 64.7]	51.1	7.2	[50.8, 51.4]
40% (T)	55.4	8.4	[55.0, 55.8]	72.3	12.2[c]	[71.8, 72.9]	55.8	7.5[a]	[55.5, 56.1]	75.2	11.0[c]	[74.7, 75.7]	50.5	6.8	[50.2, 50.8]
40% (F)	48.8	7.3	[48.5, 49.1]	47.4	7.1	[47.1, 47.7]	50.1	6.9	[49.8, 50.3]	51.4	8.6	[51.1, 51.8]	51.9	6.8	[51.6, 52.2]
50% (R)	52.6	8.4	[52.2, 52.9]	63.4	12.6[a]	[62.8, 63.9]	53.7	7.3	[53.4, 54.0]	67.3	11.0[b]	[66.9, 67.8]	51.6	6.7	[51.3, 51.9]
50% (T)	57.0	8.1[a]	[56.6, 57.3]	77.4	12.1[c]	[76.9, 77.9]	57.3	7.3[a]	[57.0, 57.7]	81.0	11.0[d]	[80.6, 81.5]	50.7	5.8	[50.4, 50.9]
50% (F)	48.4	6.7	[48.1, 48.7]	46.8	6.7	[46.5, 47.1]	50.2	5.9	[49.9, 50.4]	52.0	8.0	[51.7, 52.4]	52.3	6.1	[52.0, 52.5]
60% (R)	53.2	8.1	[52.9, 53.6]	65.6	12.4[b]	[65.0, 66.1]	54.5	7.1	[54.2, 54.8]	70.5	11.0[b]	[70.0, 70.9]	51.9	6.3	[51.7, 52.2]
60% (T)	58.1	7.5[a]	[57.8, 58.5]	82.1	11.5[d]	[81.7, 82.6]	58.9	6.8[a]	[58.6, 59.2]	86.3	10.5[e]	[85.9, 86.8]	50.8	5.0	[50.6, 51.0]
60% (F)	48.2	5.9	[48.0, 48.5]	46.2	5.7	[45.9, 46.4]	50.2	5.1	[50.0, 50.4]	52.9	7.5	[52.6, 53.3]	52.7	5.3	[52.5, 52.9]
70% (R)	53.5	8.1	[53.1, 53.8]	68.1	12.2[b]	[67.6, 68.6]	55.5	6.9[a]	[55.2, 55.8]	73.4	11.2[c]	[72.9, 73.9]	52.3	6.1	[52.0, 52.5]
70% (T)	59.9	7.1[a]	[59.6, 60.2]	86.7	10.5[e]	[86.3, 87.2]	60.3	6.0[b]	[60.1, 60.6]	92.1	9.6[f]	[91.6, 92.5]	50.8	4.2	[50.6, 50.9]
70% (F)	48.0	4.9	[47.8, 48.2]	45.7	5.1	[45.5, 45.9]	50.5	4.1	[50.3, 50.6]	53.7	6.7	[53.4, 54.0]	53.0	4.5	[52.8, 53.2]

Table 6 Continued

Response insertion percentage	STW			AXY			ANP			BRF			MSF		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
80% (R)	54.3	8.2	[53.9, 54.6]	70.4	12.0[b]	[69.9, 70.9]	55.7	6.5[a]	[55.4, 56.0]	76.2	11.3[c]	[75.7, 76.7]	52.6	5.7	[52.3, 52.8]
80% (T)	61.5	6.0[a]	[61.3, 61.8]	91.9	9.1[e]	[91.5, 92.3]	62.0	5.5[b]	[61.8, 62.2]	97.8	8.4[f]	[97.4, 98.1]	50.9	3.1	[50.7, 51.0]
80% (F)	47.6	3.8	[47.4, 47.8]	45.1	4.2	[44.9, 45.3]	50.7	3.2	[50.5, 50.8]	54.3	5.7	[54.1, 54.5]	53.5	3.7	[53.3, 53.7]
90% (R)	54.7	7.9[b]	[54.4, 55.0]	72.4	12.5[c]	[71.9, 72.9]	57.0	6.9[a]	[56.7, 57.3]	79.7	11.6[d]	[79.2, 80.2]	53.1	5.8	[52.8, 53.3]
90% (T)	63.0	4.7	[62.8, 63.2]	96.7	6.7[f]	[96.4, 96.9]	63.8	4.1[b]	[63.6, 64.0]	103.5	6.1[g]	[103.3, 103.8]	50.9	2.0	[50.8, 51.0]
90% (F)	47.6	2.6	[47.4, 47.7]	44.5	2.8	[44.4, 44.7]	50.9	2.1	[50.8, 51.0]	55.0	4.1	[54.9, 55.2]	53.8	2.6	[53.7, 53.9]
100% (R)	55.6	8.1[b]	[55.3, 56.0]	75.0	12.0[c]	[74.5, 75.5]	57.2	6.8[a]	[56.9, 57.5]	82.1	11.6[d]	[81.6, 82.6]	53.5	5.9	[53.2, 53.7]
100% (T)	65.2	0.0	[--, --]	101.5	0.0[f]	[--, --]	65.8	0.0[c]	[--, --]	108.9	0.0[h]	[--, --]	51.0	0.0	[--, --]
100% (F)	47.4	0.0	[--, --]	44.0	0.0	[--, --]	51.0	0.0	[--, --]	55.8	0.0	[--, --]	54.1	0.0	[--, --]

Note. n = 2,110 and n = 2,124 for variable and acquiescent/counter-acquiescent response insertion, respectively. RC7 =

Table 6 Continued

Dysfunctional Negative Emotions; STW = Stress/Worry; AXY = Anxiety; ANP = Anger-Proneness; BRF = Behavior-Restricting Fears; MSF = Multiple Specific Fears; R = random; T = true; and F = false.

[a,b,c,d,e,f,g,h] The mean score differs from the 0% variable, acquiescent, or counter-acquiescent baseline mean score by one, two, three, four, five, six, seven, and eight times this scale's standard error of measurement, respectively.

When fixed acquiescent and counter-acquiescent responding reaches 100%, T-scores become constant and the standard deviations equal zero.

The format of 95% confidence interval results is as follows: Mean [lower, upper bounds of 95% Confidence Interval]. [--, --] = The confidence interval for this mean score could not be calculated due to the same reason described above. The non-gendered normative sample, reported in Tellegen & Ben-Porath (2008), is the basis for all of the results presented in this table.

Table 7

The Effects of Increasing Degrees of Variable, Fixed Acquiescent, and Fixed Counter-Acquiescent Response Insertion on Normative Sample Mean Externalizing Scale T-Scores, Standard Deviations, and 95% Confidence Intervals.

Response insertion	JCP			SUB			AGG			ACT		
percentage	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
0% (R)	49.8	9.8	[49.4, 50.2]	49.8	9.9	[49.4, 50.3]	49.7	9.7	[49.3, 50.1]	49.7	9.7	[49.3, 50.1]
0% (T or F)	49.8	9.8	[49.4, 50.2]	49.8	9.9	[49.4, 50.3]	49.7	9.7	[49.2, 50.1]	49.7	9.7	[49.3, 50.1]
10% (R)	51.3	9.5	[50.9, 51.7]	51.5	9.7	[51.5, 51.9]	51.3	9.4	[50.9, 51.7]	50.0	9.3	[49.6, 50.4]
10% (T)	53.6	9.7	[53.1, 54.0]	53.4	9.6	[53.0, 53.9]	53.7	9.8	[53.3, 54.1]	52.4	9.8	[52.0, 52.8]
10% (F)	48.8	9.2	[48.4, 49.2]	49.8	9.2	[49.4, 50.2]	48.3	9.0	[47.9, 48.7]	47.7	8.9	[47.3, 48.1]
20% (R)	52.8	9.5	[52.4, 53.2]	53.1	9.3	[52.7, 53.5]	52.5	9.1	[52.2, 52.9]	50.4	9.1	[50.0, 50.8]
20% (T)	57.3	9.5[a]	[56.9, 57.7]	56.4	9.7[a]	[56.0, 56.8]	58.1	9.6[a]	[57.7, 58.5]	55.1	9.9	[54.7, 55.5]
20% (F)	48.2	8.9	[47.8, 48.6]	49.7	8.7	[49.3, 50.0]	47.5	8.5	[47.1, 47.8]	46.2	8.2	[45.9, 46.6]
30% (R)	54.2	9.0	[53.8, 54.6]	54.6	9.0	[54.2, 55.0]	54.3	9.1	[54.0, 54.7]	50.8	9.0	[50.4, 51.2]

Table 7 Continued

Response insertion percentage	JCP *M*	JCP *SD*	JCP *95% CI*	SUB *M*	SUB *SD*	SUB *95% CI*	AGG *M*	AGG *SD*	AGG *95% CI*	ACT *M*	ACT *SD*	ACT *95% CI*
30% (T)	60.6	9.4[a]	[60.2, 61.0]	59.8	10.0[a]	[59.3, 60.2]	61.9	9.7[b]	[61.5, 62.3]	58.2	10.2[a]	[57.7, 58.6]
30% (F)	46.9	8.2	[46.6, 47.3]	49.5	8.0	[49.2, 49.9]	46.2	7.9	[45.9, 46.6]	44.7	7.7	[44.3, 45.0]
40% (R)	55.5	8.8	[55.1, 55.9]	56.1	8.7[a]	[55.7, 56.5]	55.9	9.0[a]	[55.5, 56.3]	51.2	8.7	[50.8, 51.6]
40% (T)	64.1	9.1[b]	[63.7, 64.5]	63.0	10.1[b]	[62.6, 63.4]	66.3	9.4[b]	[65.9, 66.7]	61.0	10.2[a]	[60.6, 61.5]
40% (F)	46.5	7.6	[46.1, 46.8]	49.6	7.2	[49.3, 49.9]	45.2	7.3	[44.8, 45.5]	42.9	6.7[a]	[42.6, 43.2]
50% (R)	56.5	8.6[a]	[56.2, 56.9]	57.7	9.0[a]	[57.3, 58.1]	57.2	8.9[a]	[56.8, 57.6]	51.7	8.6	[51.3, 52.1]
50% (T)	67.1	8.6[b]	[66.7, 67.5]	66.2	9.9[b]	[65.8, 66.7]	69.9	9.4[c]	[69.5, 70.3]	64.6	10.1[b]	[64.2, 65.1]
50% (F)	45.3	7.0	[45.0, 45.6]	49.5	6.6	[49.2, 49.8]	43.8	6.6	[43.5, 44.1]	41.5	6.0[a]	[41.2, 41.7]
60% (R)	57.8	8.7[a]	[57.4, 58.2]	59.1	9.1[a]	[58.7, 59.5]	58.7	8.4[a]	[58.3, 59.0]	51.9	8.4	[51.5, 52.2]
60% (T)	70.3	8.1[c]	[70.0, 70.7]	69.6	9.7[c]	[69.2, 70.0]	74.6	8.5[d]	[74.2, 75.0]	68.2	9.9[b]	[67.8, 68.7]
60% (F)	44.3	6.4	[44.1, 44.6]	49.7	5.8	[49.5, 50.0]	42.8	6.2[a]	[42.6, 43.1]	39.9	5.6[a]	[49.6, 40.1]
70% (R)	59.3	8.4[a]	[58.9, 59.6]	60.5	8.9[a]	[60.1, 60.9]	60.2	8.6[a]	[59.8, 60.6]	52.7	8.4	[52.4, 53.1]
70% (T)	73.8	7.1[c]	[73.5, 74.1]	73.4	8.9[c]	[73.0, 73.8]	78.8	7.7[d]	[78.4, 79.1]	71.9	8.9[c]	[71.5, 72.3]

Table 7 Continued

Response insertion percentage	JCP *M*	JCP *SD*	JCP *95% CI*	SUB *M*	SUB *SD*	SUB *95% CI*	AGG *M*	AGG *SD*	AGG *95% CI*	ACT *M*	ACT *SD*	ACT *95% CI*
70% (F)	43.2	5.6[a]	[43.0, 43.5]	49.7	5.1	[49.5, 49.9]	41.4	5.4[a]	[41.2, 41.7]	38.3	4.8[a]	[38.1, 38.5]
80% (R)	60.8	8.4[a]	[60.5, 61.2]	62.6	9.3[a]	[62.2, 63.0]	61.7	8.6[a]	[61.3, 62.1]	53.1	8.4	[52.7, 53.4]
80% (T)	77.1	6.2[d]	[76.8, 77.4]	77.0	7.7[d]	[76.7, 77.4]	83.2	6.4[e]	[82.9, 83.4]	75.4	7.7[d]	[75.1, 75.8]
80% (F)	42.2	4.6[a]	[42.0, 42.4]	49.8	4.2	[49.6, 50.0]	40.3	4.7[a]	[40.1, 40.5]	36.6	4.1[b]	[36.5, 36.8]
90% (R)	61.9	8.3[b]	[61.5, 62.2]	63.4	9.4[a]	[63.0, 63.8]	63.0	8.7[b]	[62.6, 63.4]	53.3	8.5	[53.0, 53.7]
90% (T)	80.5	4.5[e]	[80.4, 80.7]	81.2	5.3[e]	[81.0, 81.5]	87.3	4.7[f]	[87.1, 87.5]	79.9	5.3[d]	[79.6, 80.1]
90% (F)	41.4	3.6[a]	[41.2, 41.5]	50.0	2.9	[49.9, 50.1]	38.8	3.4[a]	[38.7, 39.0]	34.8	2.8[b]	[34.6, 34.9]
100% (R)	63.2	8.2[b]	[62.8, 63.5]	65.2	9.4[a]	[64.8, 65.6]	64.6	8.7[b]	[64.2, 65.0]	54.1	8.5	[53.8, 54.5]
100% (T)	83.7	0.0[e]	[--, --]	84.8	0.0[e]	[--, --]	91.5	0.0[f]	[--, --]	83.5	0.0[e]	[--, --]
100% (F)	40.1	0.0[a]	[--, --]	50.3	0.0	[--, --]	37.3	0.0[b]	[--, --]	33.2	0.0[b]	[--, --]

Note. n = 2,110 and n = 2,124 for variable and acquiescent/counter-acquiescent response insertion,

Table 7 Continued

respectively. JCP = Juvenile

Conduct Problems; SUB = Substance Abuse; AGG = Aggression; ACT = Activation; R = random; T = true;

and F = false.

[a,b,c,d,e,f] The mean score differs from the 0% variable, acquiescent, or counter-acquiescent baseline mean score by one, two, three, four, five, and six times this scale's standard error of measurement, respectively.

When fixed acquiescent and counter-acquiescent responding reaches 100%, T-scores become constant and the standard deviations equal zero.

The format of 95% confidence interval results is as follows: Mean [lower, upper bounds of 95% Confidence Interval]. [--, --] = The confidence interval for this mean score could not be calculated due to the same reason described above. The non-gendered normative sample, reported in Tellegen & Ben-Porath (2008), is the basis for all of the results presented in this table.

Table 8

The Effects of Increasing Degrees of Variable, Fixed Acquiescent, and Fixed Counter-Acquiescent Response Insertion on Normative Sample Mean Interpersonal Scale T-Scores, Standard Deviations, and 95% Confidence Intervals.

Response insertion	FML			IPP			SAV			SHY			DSF		
percentage	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
0% (R)	49.5	9.6	[49.1, 49.9]	50.0	10.0	[49.6, 50.5]	49.8	9.9	[49.4, 50.2]	49.8	10.0	[49.4, 50.3]	49.4	9.5	[49.0, 49.8]
0% (T or F)	49.5	9.6	[49.1, 49.9]	50.1	10.0	[49.6, 50.5]	49.9	9.9	[49.5, 50.3]	49.8	10.0	[49.4, 50.3]	49.4	9.5	[49.0, 49.8]
10% (R)	50.9	9.1	[50.6, 51.3]	50.3	9.4	[50.0, 50.7]	50.7	9.1	[50.3, 51.1]	50.0	9.1	[49.6, 50.4]	53.1	10.6	[52.6, 53.5]
10% (T)	52.0	9.1	[51.6, 52.4]	48.9	9.2	[48.5, 49.3]	49.0	9.1	[48.6, 49.4]	51.2	9.3	[50.8, 51.6]	55.3	11.3	[54.8, 55.8]
10% (F)	49.9	9.0	[49.5, 50.3]	52.2	9.6	[51.8, 52.7]	52.5	9.2	[52.1, 52.9]	48.8	8.7	[48.4, 49.2]	50.5	9.5	[50.1, 50.9]
20% (R)	52.4	8.9	[52.0, 52.8]	50.7	8.7	[50.3, 51.1]	51.3	8.4	[51.0, 51.7]	50.1	8.3	[49.8, 50.5]	56.2	11.1	[55.7, 56.7]
20% (T)	54.3	8.6	[53.9, 54.6]	47.4	8.2	[47.1, 47.8]	48.3	8.2	[47.9, 48.6]	52.5	8.8	[52.2, 52.9]	60.3	12.1[a]	[59.8, 60.8]
20% (F)	50.3	8.5	[50.0, 50.7]	53.8	9.2	[53.4, 54.2]	54.3	8.6	[54.0, 54.7]	48.1	8.0	[47.7, 48.4]	51.1	9.6	[50.7, 51.5]
30% (R)	53.7	8.4	[53.3, 54.0]	51.0	8.1	[50.6, 51.3]	52.1	7.6	[51.8, 52.4]	50.6	7.6	[50.3, 50.9]	59.1	11.8[a]	[58.6, 59.6]
30% (T)	56.9	8.3[a]	[56.5, 57.2]	46.3	7.5	[46.0, 46.6]	47.5	7.5	[47.2, 47.9]	53.7	8.3	[53.3, 54.1]	65.6	12.4[b]	[65.0, 66.1]

Table 8 Continued

Response insertion percentage	FML			IPP			SAV			SHY			DSF		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
30% (F)	50.6	7.9	[50.2, 50.9]	55.9	9.0[a]	[55.5, 56.3]	56.3	8.3[a]	[56.0, 56.7]	47.3	7.1	[47.0, 47.6]	52.0	9.1	[51.7, 52.4]
40% (R)	55.1	8.3	[54.7, 55.4]	51.4	7.8	[51.1, 51.7]	52.5	7.1	[52.2, 52.8]	50.4	7.2	[50.1, 50.7]	61.6	11.9[a]	[61.1, 62.2]
40% (T)	59.0	7.8[a]	[58.7, 59.3]	45.2	6.7	[45.0, 45.5]	46.7	6.7	[46.4, 47.0]	55.1	8.1	[54.7, 55.4]	70.3	12.1[c]	[69.7, 70.8]
40% (F)	51.1	7.4	[50.8, 51.5]	57.9	8.8[a]	[57.5, 58.3]	58.4	8.1[a]	[58.1, 58.8]	46.5	6.0	[46.2, 46.7]	52.9	8.7	[52.5, 53.2]
50% (R)	56.2	8.2[a]	[55.9, 56.6]	51.4	7.5	[51.1, 51.7]	53.3	6.7	[53.0, 53.6]	50.6	6.5	[50.3, 50.9]	64.5	11.9[b]	[64.0, 65.0]
50% (T)	61.1	7.6[b]	[60.8, 61.4]	44.2	5.9[a]	[44.0, 44.5]	45.9	6.0	[45.7, 47.2]	56.3	7.7	[55.9, 56.6]	75.2	11.7[c]	[74.7, 75.7]
50% (F)	51.6	6.9	[51.4, 51.9]	60.0	8.5[a]	[59.7, 60.4]	60.9	7.8[b]	[60.5, 61.2]	46.1	5.3	[45.8, 46.3]	53.4	8.4	[53.1, 53.8]
60% (R)	57.8	8.1[a]	[57.4, 58.1]	51.8	7.0	[51.5, 52.1]	53.7	6.2	[53.5, 54.0]	51.0	6.0	[50.7, 51.2]	67.6	12.1[b]	[67.1, 68.1]
60% (T)	63.5	7.1[b]	[63.2, 63.8]	43.2	5.2[a]	[43.0, 43.4]	45.5	5.2	[45.3, 45.8]	57.5	7.2	[57.2, 57.8]	79.4	11.2[d]	[78.9, 79.9]
60% (F)	51.8	5.9	[51.6, 52.1]	62.5	8.0[b]	[62.2, 62.9]	63.3	7.4[b]	[63.1, 63.7]	45.4	4.6	[45.2, 45.6]	54.7	7.9	[54.3, 55.0]
70% (R)	59.0	7.9[a]	[58.6, 59.3]	52.4	6.7	[52.1, 52.6]	54.2	6.0	[53.9, 54.4]	51.2	5.6	[50.9, 51.4]	69.8	12.4[b]	[69.3, 70.3]
70% (T)	66.1	6.4[b]	[65.8, 66.4]	42.1	4.3[a]	[41.9, 42.3]	44.8	4.6[a]	[44.6, 45.0]	59.2	6.6	[59.0, 59.5]	84.5	9.9[e]	[84.1, 85.0]
70% (F)	52.3	5.1	[52.1, 52.6]	65.5	7.3[b]	[65.2, 65.8]	66.0	6.8[c]	[65.7, 66.3]	45.1	4.0	[45.0, 45.3]	55.8	7.0[c]	[55.5, 56.1]

Table 8 Continued

Response insertion percentage	FML *M*	FML *SD*	FML *95% CI*	IPP *M*	IPP *SD*	IPP *95% CI*	SAV *M*	SAV *SD*	SAV *95% CI*	SHY *M*	SHY *SD*	SHY *95% CI*	DSF *M*	DSF *SD*	DSF *95% CI*
80% (R)	60.2	7.8[a]	[59.8, 60.5]	52.5	6.7	[52.2, 52.8]	55.1	6.0[a]	[54.8, 55.3]	51.4	5.4	[51.1, 51.6]	73.2	12.3[c]	[72.7, 73.8]
80% (T)	68.6	5.5[c]	[68.3, 68.8]	41.1	3.7[a]	[40.9, 41.2]	44.2	3.9[a]	[44.0, 44.3]	61.1	6.1	[60.8, 61.3]	88.9	8.4[e]	[88.6, 89.3]
80% (F)	52.7	4.1	[52.5, 52.9]	68.4	6.3[c]	[68.1, 68.6]	69.0	5.8[d]	[68.8, 69.3]	44.6	3.4[a]	[44.5, 44.8]	56.3	6.0	[56.1, 56.6]
90% (R)	62.2	7.8[b]	[61.8, 62.5]	53.3	6.4	[53.0, 53.6]	56.0	5.9[a]	[55.7, 56.2]	51.5	5.2	[51.3, 51.7]	75.7	12.2[c]	[75.2, 76.2]
90% (T)	71.2	3.8[c]	[71.0, 71.3]	40.2	2.8[a]	[40.1, 40.3]	43.5	2.7[a]	[43.4, 43.6]	63.2	4.7	[63.0, 63.4]	93.6	6.3[f]	[93.4, 93.9]
90% (F)	52.9	2.8	[52.8, 53.1]	71.3	4.6[c]	[71.1, 71.5]	72.2	4.1[d]	[72.0, 72.4]	44.2	2.5[a]	[44.1, 44.3]	57.5	4.2[a]	[57.3, 57.7]
100% (R)	62.9	8.0[b]	[62.5, 63.2]	53.6	6.5	[53.3, 53.8]	56.2	5.9[a]	[55.9, 56.4]	51.5	5.1	[51.3, 51.7]	77.4	12.4[d]	[76.8, 77.9]
100% (T)	73.6	0.0[d]	[--, --]	39.1	0.0[b]	[--, --]	42.9	0.0[a]	[--, --]	65.7	0.0	[--, --]	97.8	0.0[f]	[--, --]
100% (F)	53.3	0.0	[--, --]	74.4	0.0[d]	[--, --]	75.2	0.0[e]	[--, --]	43.8	0.0[a]	[--, --]	58.3	0.0[a]	[--, --]

Note. n = 2,110 and n = 2,124 for variable and acquiescent/counter-acquiescent response insertion, respectively. FML = Family Problems; IPP = Interpersonal Passivity; SAV = Social Avoidance; SHY = Shyness; DSF = Disaffiliativeness; R = random;

Table 8 Continued

T = true; F = false.

[a,b,c,d,e,f] The mean score differs from the 0% variable, acquiescent, or counter-acquiescent baseline mean score by one, two, three, four, five, and six times this scale's standard error of measurement, respectively.

When fixed acquiescent and counter-acquiescent responding reaches 100%, T-scores become constant and the standard deviations equal zero.

The format of 95% confidence interval results is as follows: Mean [lower, upper bounds of 95% Confidence Interval]. [--, --] = The confidence interval for this mean score could not be calculated due to the same reason described above. The non-gendered normative sample, reported in Tellegen & Ben-Porath (2008), is the basis for all of the results presented in this table.

Table 9

The Effects of Increasing Degrees of Variable, Fixed Acquiescent, and Fixed Counter-Acquiescent Response Insertion on Normative Sample Mean PSY-5 Scale T-Scores, Standard Deviations, and 95% Confidence Intervals.

Response insertion	AGGR-r			PSYC-r			DISC-r			NEGE-r			INTR-r		
percentage	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
0% (R)	49.8	9.9	[49.4, 50.3]	49.2	9.2	[48.8, 49.6]	49.9	9.9	[49.5, 50.3]	49.5	9.5	[49.1, 49.9]	49.8	9.7	[49.4, 50.2]
0% (T or F)	49.8	9.9	[49.4, 50.2]	49.2	9.3	[48.8, 49.6]	49.9	9.9	[49.5, 50.3]	49.5	9.5	[49.1, 49.9]	49.9	9.7	[49.5, 50.3]
10% (R)	49.9	9.2	[49.5, 50.3]	54.7	8.4[a]	[54.4, 55.1]	50.9	9.2	[50.5, 51.3]	50.5	8.9	[50.1, 50.9]	51.0	9.1	[50.6, 51.3]
10% (T)	51.8	9.9	[51.4, 52.3]	59.5	8.4	[59.1, 59.8]	52.8	9.4	[52.4, 53.2]	51.8	9.1	[51.4, 52.2]	48.2	8.9	[47.8, 48.5]
10% (F)	47.7	8.6	[47.3, 48.0]	48.9	8.8	[48.5, 49.3]	48.9	8.8	[48.5, 49.2]	49.1	8.5	[48.8, 49.5]	54.1	9.5	[53.7, 54.5]
20% (R)	49.7	8.5	[49.3, 50.0]	59.3	8.1[c]	[58.9, 59.6]	51.8	8.8[a]	[51.4, 52.2]	51.4	8.3	[51.0, 51.7]	52.0	8.6	[51.6, 52.4]
20% (T)	54.1	9.8	[53.7, 54.5]	67.7	8.3	[67.4, 68.1]	55.7	9.0	[55.3, 56.1]	54.1	8.9	[53.7, 54.5]	46.5	8.1[a]	[46.2, 46.9]
20% (F)	46.1	7.4	[45.8, 46.4]	48.8	8.3	[48.4, 49.1]	48.1	8.3	[47.7, 48.5]	48.9	7.7	[48.5, 49.2]	57.6	9.3	[57.3, 58.0]
30% (R)	49.9	8.1[a]	[49.6, 50.3]	63.6	7.7[d]	[63.2, 63.9]	52.8	8.0[a]	[52.4, 53.1]	52.2	7.9[a]	[51.9, 52.5]	53.0	8.0	[52.6, 53.3]
30% (T)	56.5	9.9[a]	[56.1, 56.9]	75.7	8.6	[75.3, 76.1]	58.7	8.6	[58.4, 59.1]	56.2	8.7	[55.8, 56.6]	45.0	7.3[b]	[44.7, 45.3]

Table 9 Continued

Response insertion percentage	AGGR-r			PSYC-r			DISC-r			NEGE-r			INTR-r		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
30% (F)	44.3	6.4	[44.1, 44.6]	48.3	7.8	[48.0, 48.6]	47.1	7.3	[46.8, 47.4]	48.6	7.0	[48.3, 48.9]	61.6	9.1	[61.2, 62.0]
40% (R)	49.8	7.5[a]	[49.5, 50.2]	67.3	8.0[f]	[67.0, 67.6]	53.8	7.8[b]	[53.5, 54.2]	53.3	7.4[a]	[52.9, 53.6]	54.1	7.7[a]	[53.8, 54.4]
40% (T)	59.2	9.7[a]	[58.8, 59.6]	83.7	8.8	[83.3, 84.0]	61.7	8.2	[61.4, 62.1]	58.8	8.3	[58.5, 59.2]	43.1	6.6[c]	[42.8, 43.4]
40% (F)	43.0	5.5	[42.7, 43.2]	48.0	7.2	[47.7, 48.3]	46.3	6.5	[46.1, 46.1]	48.3	6.2	[48.0, 48.5]	66.1	8.7	[65.8, 66.5]
50% (R)	50.2	7.3[b]	[49.9, 50.5]	71.0	8.1[g]	[70.7, 71.4]	54.6	7.3[c]	[54.3, 55.0]	54.1	7.1[b]	[53.8, 54.5]	54.9	7.3[a]	[54.6, 55.2]
50% (T)	61.7	9.2[a]	[61.3, 62.1]	90.8	8.3	[90.5, 91.2]	64.9	7.8	[64.6, 65.2]	61.5	8.2	[61.1, 61.8]	41.4	5.8[d]	[41.2, 41.7]
50% (F)	41.6	4.8	[41.4, 41.8]	48.1	6.8	[47.8, 48.4]	45.5	5.9	[45.2, 45.7]	47.9	5.3	[47.7, 48.2]	70.5	8.1	[70.2, 70.9]
60% (R)	50.3	6.9[b]	[50.0, 50.6]	74.5	8.1[h]	[74.1, 74.8]	55.6	6.9[c]	[55.3, 55.9]	55.2	6.9[c]	[54.9, 55.5]	56.1	7.0[b]	[55.8, 56.4]
60% (T)	65.0	8.8[a]	[64.6, 65.4]	98.7	8.1	[98.4, 99.0]	68.1	7.2[a]	[67.8, 68.4]	64.1	7.5	[63.7, 64.4]	39.8	5.2[d]	[39.6, 40.1]
60% (F)	40.3	4.0	[40.1, 40.5]	47.6	6.0	[47.3, 47.8]	44.7	5.0	[44.5, 44.9]	47.7	4.4	[47.5, 47.9]	75.1	7.4	[74.7, 75.4]
70% (R)	50.2	6.5[c]	[49.9, 50.4]	78.5	8.3[i]	[78.1, 78.8]	56.6	6.7[d]	[56.3, 56.9]	55.9	6.7[c]	[55.7, 56.2]	57.1	6.9[b]	[56.8, 57.4]
70% (T)	68.2	7.7[b]	[67.9, 68.6]	106.3	7.3	[106.0, 106.6]	71.6	6.1[a]	[71.3, 71.8]	67.2	6.8	[66.9, 67.4]	38.1	4.5[e]	[37.9, 38.3]
70% (F)	38.8	3.4	[38.6, 38.9]	47.3	5.2	[47.1, 47.6]	43.8	4.3	[43.6, 44.0]	47.6	3.5	[47.4, 47.7]	79.5	6.6	[79.2, 79.7]

Table 9 Continued

Response insertion percentage	AGGR-r *M*	AGGR-r *SD*	AGGR-r *95% CI*	PSYC-r *M*	PSYC-r *SD*	PSYC-r *95% CI*	DISC-r *M*	DISC-r *SD*	DISC-r *95% CI*	NEGE-r *M*	NEGE-r *SD*	NEGE-r *95% CI*	INTR-r *M*	INTR-r *SD*	INTR-r *95% CI*
80% (R)	50.4	6.4[d]	[50.1, 50.6]	81.9	8.6[i]	[81.6, 82.3]	57.8	6.5[e]	[57.5, 58.1]	57.1	6.8[d]	[56.8, 57.4]	58.5	6.8[b]	[58.2, 58.8]
80% (T)	71.9	6.5[b]	[71.6, 72.2]	114.4	6.5	[114.1, 114.7]	75.0	5.2[a]	[74.8, 75.3]	70.1	5.6	[69.9, 70.3]	36.4	3.8[f]	[36.2, 36.5]
80% (F)	37.4	2.9	[37.3, 37.6]	47.0	4.5	[46.8, 47.2]	42.9	3.4	[42.7, 43.0]	47.3	2.7	[47.2, 47.4]	84.3	5.4	[84.1, 84.5]
90% (R)	50.0	6.0[d]	[49.8, 50.3]	86.1	8.4[i]	[85.7, 86.4]	58.5	6.4[e]	[58.2, 58.8]	58.2	6.8[e]	[57.9, 58.5]	59.9	6.9[c]	[59.6, 60.2]
90% (T)	75.1	4.8[b]	[74.9, 75.3]	122.2	4.6	[122.0, 122.4]	78.6	3.6[a]	[78.5, 78.8]	73.4	4.0	[73.2, 73.6]	34.3	2.7[g]	[34.1, 34.4]
90% (F)	36.1	2.1	[36.1, 36.2]	47.0	3.1	[47.9, 47.1]	42.2	2.3	[42.1, 42.3]	47.1	1.9	[47.1, 47.2]	88.8	3.8	[88.7, 89.0]
100% (R)	50.4	6.1[e]	[50.1, 50.6]	89.5	8.5[i]	[89.1, 89.9]	59.7	6.6[f]	[59.4, 60.0]	59.4	7.0[e]	[59.1, 59.7]	60.5	6.9[c]	[60.2, 60.8]
100% (T)	78.5	0.0[b]	[--, --]	129.7	0.0	[--, --]	81.9	0.0[a]	[--, --]	76.7	0.0	[--, --]	32.2	0.0[h]	[--, --]
100% (F)	34.6	0.0	[--, --]	46.6	0.0	[--, --]	41.5	0.0	[--, --]	46.9	0.0	[--, --]	93.3	0.0	[--, --]

Note. n = 2,110 and n = 2,124 for variable and acquiescent/counter-acquiescent response insertion, respectively. PSY-5 = Personality Psychopathology Five; AGGR-r = Aggressiveness-Revised; PSYC-r = Psychoticism-Revised; DISC-r = Disconstraint-Revised;

Table 9 Continued

NEGE-r = Negative Emotionality/Neuroticism-Revised; INTR-r = Introversion/Low Positive Emotions-Revised; R = random; T = true; and F = false.

[a,b,c,d,e,f,g,h,i] The mean score differs from the 0% variable, acquiescent, or counter-acquiescent baseline mean score by one, two, three, four, five, six, seven, eight, and nine or more times this scale's standard error of measurement, respectively.

When fixed acquiescent and counter-acquiescent responding reaches 100%, T-scores become constant and the standard deviations equal zero.

The format of 95% confidence interval results is as follows: Mean [lower, upper bounds of 95% Confidence Interval]. [--, --] = The confidence interval for this mean score could not be calculated due to the same reason described above. The non-gendered normative sample, reported in Tellegen & Ben-Porath (2008), is the basis for all of the results presented in this table.

For the random response insertion analyses, a change of ≥ 5 T-score points was observed for the following SP and PSY-5 Scales. Somatic/Cognitive Scale scores increased by ≥ 5 T-score points at random response insertion rates ranging from 20% (GIC and NUC) to 40% (MLS). Internalizing Scale T-scores increased at rates ranging from 20% (SUI, AXY, and BRF) to 100% (STW). Externalizing Scale T-scores increased at rates ranging from 40% (SUB) and, for the ACT Scale, did not change by ≥ 5 T-score points at 100% random response insertion. Interpersonal Scale T-scores increased at rates ranging from 20% (DSF) and, for the IPP and SHY Scales, did not change by ≥ 5 T-score points at 100% random response insertion. PSY-5 Scale T-scores increased at rates ranging from 20% (PSYC-r) and, for the AGGR-r Scale, did not change by ≥ 5 T-score points at 100% random response insertion. Therefore, NUC, GIC, SUI, AXY, BRF, DSF, and PSYC-r were the scales most susceptible to random response insertion, as evidenced by mean scale changes of ≥ 5 T-score points at 20% simulated random response insertion. It should be noted that at random response insertion rates of 20%, the VRIN-r mean score was less than 80T. This indicates that VRIN-r would not identify a substantial portion of these cases as invalid. Specifically, Handel et al. (2010) reported that 8.1% of cases reached a VRIN-r T-score of ≥ 80T at 20% random response insertion.

In addition to examining SP and PSY-5 Scale mean T-scores, 95% confidence intervals and SEM values were calculated. Confidence intervals between levels of response insertion described above (e.g., 20% random response insertion for SUI) and baseline (i.e., 0% response insertion) were non-overlapping. Further, the confidence intervals for mean T-scores at 100% random response insertion had limited overlap as compared to those for the baseline conditions. However, as previously noted, mean T-

scores can still be statistically significantly different from one another even in the presence of overlapping confidence intervals (Wolfe & Hanley, 2002).

SEM values spanned from 5 T-score points (SAV, SHY, DISC-r, NEGE-r, and INTR-r) to 8 T-score points (SUI) for the 28 SP and PSY-5 Scales. The magnitude of deviation between mean scale T-scores at 100% response insertion and baseline ranged from 0 to 9 SEMs. A change by an SEM of at least one was observed for scales in each SP and PSY-5 Scale cluster at the following levels of random response insertion: 20% (NUC) to 50% (MLS) for the Somatic/Cognitive Scales; 30% (SUI) to 70% (NFC) for the RCd-Associated Internalizing SP Scales; 20% (BRF) and the MSF scale did not deviate by one SEM for the RC-7 Internalizing Scales; 40% (SUB) and the ACT scale did not deviate by one SEM for the Externalizing Scales; 30% (DSF) and the IPP and SHY Scales did not deviate by one SEM for the Interpersonal Scales; and 10% (PSYC-r) to 40% (INTR-r) for the PSY-5 Scales. Therefore, the PSYC-r and GIC, BRF, and DISC-r Scales were the most susceptible to random response insertion at 10% and 20% insertion, respectively. As noted above, because the mean VRIN-r T-Score at these levels was less than 80T, VRIN-r would not identify a portion of these cases.

Fixed response insertion. The results of increasing degrees of simulated fixed acquiescent and counter-acquiescent responding on SP and PSY-5 Scale T-scores are also presented in Tables 1 through 6. Under conditions of simulated increasing degrees of fixed acquiescent (i.e., “true”) responding, 24 of the 28 SP and PSY-5 Scales increased monotonically. The remaining four scales, which included MLS, IPP, SAV, and INTR-r, decreased monotonically. Under conditions of simulated increasing degrees of fixed counter-acquiescent (i.e., “false”) responding, 18 of the 28 SP Scales increased

monotonically. The remaining 10 scales, which included SUI, SFD, NFC, STW, AXY, SHY, AGGR-r, PSYC-r, DISC-r, and NEGE-r decreased monotonically. As with the VRIN-r analyses, however, SP and PSY-5 Scale T-scores differed in their susceptibility to fixed acquiescent and counter-acquiescent responding. These differences are discussed below, using the same three indicators presented in the VRIN-r results.

Acquiescent response insertion. For the fixed acquiescent response insertion analyses, a change of ≥ 5 T-score points was observed for the following SP and PSY-5 Scales. Somatic/Cognitive Scale T-scores increased by ≥ 5 T-score points at insertion rates ranging from 10% (GIC) to 40% (NUC); a decrease of ≥ 5 T-score points at 100% response insertion for MLS was not observed. Internalizing Scale T-scores increased at rates ranging from 10% (SUI, AXY, and BRF) and, for the MSF scale, did not change by ≥ 5 T-score points. Each of the four Externalizing Scale T-scores (JCP, SUB, AGG, and ACT) increased by ≥ 5 T-score points at 20% response insertion. Interpersonal Scale T-scores increased at rates ranging from 10% (DSF) to 40% (SHY); a decrease of ≥ 5 T-score points was observed at 50% and 70% response insertion for IPP and SAV, respectively. PSY-5 Scale T-scores increased by ≥ 5 T-score points at response rates ranging from 10% (PSYC-r) to 30% (NEGE-r); a decrease of ≥ 5 T-score points was observed at 40% for INTR-r. Therefore, GIC, SUI, AXY, BRF, DSF, and PSYC-r were the scales most susceptible to fixed acquiescent response insertion, as evidenced by mean scale changes of ≥ 5 T-score points at 10% simulated acquiescent response insertion. At 20% acquiescent response insertion, the COG, HLP, NFC, JCP, SUB, AGG, ACT, and DISC-r Scales evidenced a T-score change of this magnitude. As with the random response insertion results presented previously, the TRIN-r mean scores at 10% and 20%

fixed acquiescent response insertion were less than 80T. Therefore, TRIN-r would not identify a portion of these cases as invalid. Specifically, Handel et al. (2010) reported that 29.4% of cases reached a TRIN-r T-score of ≥ 80T with 20% acquiescent response insertion.

In addition to examining SP and PSY-5 Scale mean T-scores, 95% confidence intervals and SEM values were calculated. Confidence intervals indicated that the differences in mean T-scores between levels of response insertion described above (e.g., 10% random response insertion for GIC) and baseline (i.e., 0% response insertion) were non-overlapping. Further, the confidence intervals for mean T-scores at 90% fixed acquiescent response insertion had limited overlap as compared to those for the baseline conditions. However, as discussed previously, mean T-scores can still be significantly different from one another even in the presence of overlapping confidence intervals (Wolfe & Hanley, 2002). Confidence intervals at 100% response insertion could not be calculated because the standard deviation becomes zero when all responses become constant (i.e., all items are *true* responses).

Regarding the SEM analyses, SEM values spanned from 5 T-score points (SAV, SHY, DISC-r, NEGE-r, and INTR-r) to 8 T-score points (SUI) for the 28 SP and PSY-5 Scales. The magnitude of deviation between mean scale T-scores at 100% response insertion and baseline ranged from 0 to 14. A change by an SEM of at least one was observed for scales in each SP and PSY-5 Scale cluster at the following levels of fixed acquiescent response insertion: 10% (GIC) and the MLS scale did not deviate by one SEM for the Somatic/Cognitive Scales; 10% (SUI) to 30% (HLP and SFD) for the RCd-Associated Internalizing Scales; 20% (BRF) and the MSF scale did not deviate by one

SEM for the RC-7 Internalizing Scales; 20% (JCP, SUB, and AGG) to 30% (ACT) for the Externalizing Scales; 20% (DSF) and the SHY scale did not deviate by one SEM for the Interpersonal Scales; and 20% (INTR-r) and the PSYC-r and NEGE-r Scales did not deviate by one SEM for the PSY-5 Scales. Therefore, the GIC, SUI, BRF, and PSYC-r Scales were the most susceptible to fixed acquiescent response insertion, as evidenced by a change of one SEM at 10% insertion. At 20%, the COG, NFC, AXY, JCP, SUB, AGG, DSF, and DISC-r Scales evidenced a change of this magnitude. As noted above, mean TRIN-r T-scores at these levels of response insertion were less than 80T. Therefore, TRIN-r would not identify a portion of these cases.

Counter-acquiescent response insertion. For the fixed counter-acquiescent response insertion analyses, a change of ≥ 5 T-points was observed for the following SP and PSY-5 Scales. Somatic/Cognitive Scale T-scores increased by ≥ 5 T-score points at insertion rates ranging from 10% (NUC) to 100% (COG). Internalizing Scale T-scores increased at rates ranging from 90% (BRF) and did not change by ≥ 5 T-score points for HLP, ANP, or MSF; a decrease of ≥ 5 T-score points was observed 40% (NFC) and did not change by this amount for SUI and STW. Each of the four Externalizing Scale T-scores (JCP, SUB, AGG, and ACT) increased by ≥ 5 T-score points at 20% response insertion. Interpersonal Scale T-scores increased at rates ranging from 30% (IPP) and, for the FML scale, did not change by ≥ 5 T-score points; a decrease of ≥ 5 T-score points was observed at 80% for SHY. PSY-5 Scale T-scores decreased at rates ranging from 40% (AGGR-r) and did not change by ≥ 5 T-score points for PSYC-r and NEGE-r; an increase of ≥ 5 T-score points was observed at 20% for INTR-r. Therefore, the NUC Scale was the most susceptible to fixed counter-acquiescent response insertion, as evidenced by a mean

scale change of ≥ 5 T-score points at 10% simulated response insertion. At 20% response insertion, the MLS and INTR-r Scales evidenced mean T-score changes of this magnitude. As with previous results, the TRIN-r mean scores at 10% and 20% fixed counter-acquiescent response insertion were less than 80T. Therefore, TRIN-r would likely not identify a portion of these cases as invalid. Specifically, Handel et al. (2010) reported that 16.5% of cases reached a TRIN-r T-score of ≥ 80T at 20% response insertion.

In addition to examining SP and PSY-5 Scale mean T-scores, 95% confidence intervals and SEM values were calculated. Confidence intervals for mean T-scores between levels of response insertion described above (e.g., 10% fixed counter-acquiescent response insertion for NUC) and baseline (i.e., 0% response insertion) were non-overlapping. Further, the confidence intervals for mean T-scores at 90% fixed counter-acquiescent response insertion had limited overlap as compared to baseline conditions. However, as discussed previously, mean T-scores can still be significantly different from one another even in the presence of overlapping confidence intervals (Wolfe & Hanley, 2002). Confidence intervals at 100% response insertion could not be calculated because the standard deviation becomes zero when all responses become constant (i.e., all items were changed to *false* responses).

Regarding the SEM analyses, SEM values spanned from 5 T-score points (SAV, SHY, DISC-r, NEGE-r, and INTR-r) to 8 T-score points (SUI) for the 28 SP and PSY-5 Scales. The magnitude of deviation between mean scale T-scores at 100% response insertion and baseline conditions ranged from 0 to 9. A change by an SEM of at least one was observed for scales in each SP and PSY-5 Scale cluster at the following levels of

counter-acquiescent response insertion: 20% (MLS and NUC) and the COG scale did not deviate by one SEM for the Somatic/Cognitive Scales; 50% (NFC) and the SUI and HLP scales did not deviate by one SEM for the RCd-Associated Internalizing Scales; none of the RC-7 Internalizing Scales deviated by one SEM; 40% (ACT) and the SUB scale did not deviate by one SEM for the Externalizing Scales; 30% (IPP and SAV) and the FML Scale did not deviate by one SEM for the Interpersonal Scales; and 20% (INTR-r) and the PSYC-r and NEGE-r scales did not deviate by one SEM for the PSY-5 Scales. Therefore, the MLS, NUC, and INTR-r Scales were the most susceptible to fixed counter-acquiescent response insertion, as evidenced by a change of one SEM at 20% response insertion. As noted above, because the mean TRIN-r T-Score at this level of response insertion was less than 80T, TRIN-r would not identify a portion of these cases.

Psychiatric Inpatient Sample Analyses

Results from the inpatient sample analyses are presented in two primary clusters. In examination of Hypothesis 1, the effects of simulated random and fixed responding on SP and PSY-5 scales will be presented. In examination of Hypotheses 2a and 2b, results of the validity analyses will be reviewed.

As with the normative sample analyses, very conservative validity criteria were applied prior to beginning the random, fixed acquiescent, and fixed counter-acquiescent insertion analyses for the psychiatric inpatient sample. For the random insertion analyses, the following criteria used were: CNS < 15; TRIN-r < 80; F-r < 90; Fp-r < 80; L-r < 80; RBS < 80; FBS-r < 80; and Fs < 80 (Ben-Porath, 2012). For the fixed insertion analyses, the following criteria were used: CNS < 15; VRIN-r < 80; F-r < 90; Fp-r < 80; L-r < 80; RBS < 80; FBS-r < 80; and Fs < 80. These conservative criteria were applied in order to

examine the effects of simulated random and fixed responding in the absence of the possible confounding effects of even moderate levels of other invalidating response styles (i.e., overreporting or underreporting). Table 4 displays the number of protocols eliminated from the total inpatient sample (N = 704) by a sequential application of individual validity criteria. Table 5 illustrates the frequency, as represented by a numerical count and percentage, of protocols from the total normative sample that would be identified as invalid by individual validity criteria. It should be noted that the number of protocols identified in Table 5 exceeded the number of protocols excluded by the sequential application of validity criteria represented in Table 4. This is due to the fact that eliminating protocols sequentially (e.g., VRIN-r, then F-r) reduces the number of protocols that can be eliminated subsequently. Application of these criteria resulted in sample sizes of n = 277 and n = 275 for VRIN-r and TRIN-r, respectively. Frequency counts of each validity measure were calculated after the application of these criteria to confirm that there were not any invalid protocols remaining in the two samples to be used in the insertion analyses.

Table 10

Protocol Elimination From the Inpatient Sample (N = 704) as a Function of Sequential Application of MMPI-2-RF Validity Criteria

Validity Scale Name	Random Insertion Condition #	Fixed Insertion Condition #
CNS	28	28
VRIN-r	--	53
TRIN-r	88	--
F-r	210	242
Fp-r	20	24
Fs	37	39
FBS-r	19	19
RBS	11	11
L-r	14	13
Total	427	429
Resulting Sample Size	277	275

Note: # = number of protocols eliminated by each validity criteria; CNS = Cannot Say; VRIN-r = Variable Response Inconsistency – Revised; TRIN-r = True Response Inconsistency – Revised; F-r = Infrequent Responses – Revised; Fp-r = Infrequent Psychopathology Responses – Revised; Fs = Infrequent Somatic Responses; FBS-r =

Table 10 Continued

Symptom Validity – Revised; RBS = Response Bias; L-r = Uncommon Virtues – Revised. [--] = The validity criterion was not applied to this condition.

Table 11

The Frequency of Invalid Protocols in the Inpatient Sample (N = 704) as Identified by MMPI-2-RF Validity Criteria

Validity Scale Name	Protocols Identified #	(%)
CNS	28	(0.040)
VRIN-r	56	(0.080)
TRIN-r	88	(0.125)
F-r	290	(0.412)
Fp-r	194	(0.276)
Fs	254	(0.361)
FBS-r	156	(0.011)
RBS	237	(0.011)
L-r	23	(0.032)

Note: # = number of protocols eliminated by each validity criteria; (%) = percentage of the total number of protocols identified as invalid by each validity criteria; CNS = Cannot Say; VRIN-r = Variable Response Inconsistency – Revised; TRIN-r = True Response Inconsistency – Revised; F-r = Infrequent Responses – Revised; Fp-r = Infrequent Psychopathology Responses – Revised; Fs = Infrequent Somatic Responses; FBS-r =

Table 11 Continued

Symptom Validity – Revised; RBS = Response Bias; L-r = Uncommon Virtues – Revised.

A Pearson chi-square analysis was conducted to determine if the rate of invalid protocol identification was statistically significantly different for protocols completed by Caucasian versus African American participants. Overall, protocols from African American participants were identified as invalid at a small, but statistically significant higher rate than protocols from Caucasian participants (χ^2 (1) = 14.100, $p \leq .001$, φ = .153). Of 141 total African American participants, 106 (75.2%) were identified as invalid; 265 (57.6%) of 460 total Caucasian participants were identified as invalid. These high rates of invalid protocols were due to the fact that extremely conservative validity criteria were used in the present study (i.e., criteria that were much more conservative than those that would be used in actual clinical practice). Further analyses revealed that the VRIN-r mean T-score for Caucasian participants was significantly lower than the mean T-score for African American participants ($t(599) = -5.433$, $p < .001$, Cohen's $d = -0.509$); a significant racial difference was also found for the TRIN-r mean T-score, such that Caucasian participants had significantly lower mean T-scores than African American participants ($t(599) = -5.814$, $p < .001$, Cohen's $d = -0.493$). Participants from both racial categories had highly similar levels of education (Caucasian: $M = 11.89$, $SD = 2.3$; African American: $M = 11.93$, $SD = 2.2$); other variables (e.g., reading level) that could have been used to further investigate this difference were not available in this data set. It should be noted that these differences do not have a bearing on subsequent insertion analyses.

Random response insertion. The results of increasing degrees of simulated random responding on inpatient SP and PSY-5 Scale mean T-scores are presented in Tables 7 through 12.

Table 12

The Effects of Increasing Degrees of Variable, Fixed Acquiescent, and Fixed Counter-Acquiescent Response Insertion on Mean Inpatient Sample Somatic/Cognitive Scale T-Scores, Standard Deviations, and 95% Confidence Intervals.

Response insertion	MLS			GIC			HPC			NUC			COG		
percentage	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
0% (R)	55.6	11.3	[54.3, 57.0]	52.5	11.1	[51.2, 53.8]	51.7	11.2	[50.4, 53.0]	53.9	10.7	[52.6, 55.2]	51.2	11.1	[49.8, 52.3]
0% (T or F)	55.8	11.3	[54.5, 57.2]	52.7	11.1	[51.4, 54.0]	52.0	11.3	[50.7, 53.4]	53.9	10.5	[52.6, 55.1]	51.1	10.6	[49.8, 53.4]
10% (R)	56.6	10.6	[55.3, 57.9]	56.5	11.7	[55.1, 57.9]	53.6	11.1	[52.3, 54.9]	56.9	10.5	[55.6, 58.1]	53.1	9.9	[52.0, 54.3]
10% (T)	54.8	10.8	[53.5, 56.0]	57.8	12.1	[56.3, 59.2]	53.5	10.9	[52.2, 54.8]	54.9	10.2	[53.6, 56.1]	54.6	10.0	[53.4, 55.8]
10% (F)	58.1	10.3	[56.9, 59.3]	53.7	11.0	[52.4, 55.0]	53.1	10.0	[51.8, 54.4]	57.7	9.6	[56.6, 58.9]	51.2	9.7	[50.1, 52.4]
20% (R)	57.6	9.8	[56.4, 58.7]	58.5	12.0	[57.1, 59.9]	55.2	10.4	[54.0, 56.4]	59.4	9.6	[58.3, 60.6]	55.6	8.9	[54.5, 56.6]
20% (T)	53.4	9.6	[52.3, 54.5]	62.7	12.4[a]	[61.2, 64.1]	55.2	9.8	[54.1, 56.4]	56.8	9.6	[55.7, 57.9]	58.5	9.5[a]	[57.4, 59.6]
20% (F)	60.8	10.3	[59.6, 62.0]	54.9	11.4	[53.6, 56.3]	55.2	9.8	[54.0, 56.3]	62.3	9.9[a]	[61.1, 63.5]	51.9	9.3	[50.8, 53.0]
30% (R)	58.2	9.8	[57.1, 59.4]	61.3	12.1[a]	[59.9, 62.7]	56.0	9.3	[54.9, 57.1]	61.9	9.3[a]	[60.8, 63.0]	56.6	9.2	[55.5, 57.7]
30% (T)	52.6	9.5	[51.5, 53.7]	66.7	10.8[a]	[65.4, 68.0]	56.6	9.6	[55.4, 57.7]	57.5	8.8	[56.4, 58.5]	62.1	9.2[a]	[61.0, 63.2]

Table 12 Continued

Response insertion percentage	MLS			GIC			HPC			NUC			COG		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
30% (F)	62.9	9.0[a]	[61.8, 64.0]	55.9	11.0	[54.6, 57.2]	56.6	9.1	[55.6, 57.7]	65.3	9.6[a]	[64.2, 66.4]	52.2	8.4	[51.2, 53.2]
40% (R)	59.0	9.2	[57.9, 60.1]	62.9	11.9[a]	[61.5, 64.3]	57.9	9.1[a]	[56.8, 58.9]	63.8	9.5[a]	[62.7, 64.9]	58.2	8.2	[57.3, 59.2]
40% (T)	51.4	8.8	[50.3, 52.4]	70.9	8.9[b]	[69.8, 71.9]	57.8	8.6	[56.8, 58.8]	58.8	8.2	[57.8, 59.7]	64.6	8.9[b]	[63.5, 65.6]
40% (F)	65.2	8.5[a]	[64.1, 66.2]	56.5	10.7	[55.3, 57.8]	57.5	8.5	[56.5, 58.5]	67.8	9.1[a]	[66.7, 68.8]	52.6	7.5	[51.7, 53.5]
50% (R)	59.2	8.6	[58.2, 60.2]	65.5	11.0[a]	[64.2, 66.8]	58.1	9.3[a]	[56.9, 59.2]	65.2	9.3[a]	[64.1, 66.3]	60.7	8.2[a]	[59.8, 61.7]
50% (T)	51.0	8.1	[50.0, 51.9]	74.4	8.3[b]	[73.4, 75.4]	59.4	7.6[a]	[58.5, 60.3]	59.7	7.9	[58.7, 60.6]	67.9	8.6[b]	[66.9, 68.9]
50% (F)	67.8	8.0[a]	[66.9, 68.8]	58.2	10.0	[57.0, 59.4]	58.4	8.5[a]	[57.4, 59.4]	71.1	7.7[b]	[70.1, 72.0]	53.0	7.2	[52.2, 53.9]
60% (R)	60.6	8.3	[59.6, 61.6]	68.3	10.4[a]	[67.0, 69.5]	60.4	8.8[a]	[59.4, 61.5]	68.5	8.6[a]	[67.5, 69.5]	62.8	7.9[a]	[61.9, 63.7]
60% (T)	50.1	6.9	[49.3, 51.0]	76.5	9.4[b]	[75.4, 77.6]	60.6	6.7[a]	[59.8, 61.4]	60.9	6.8	[60.1, 61.7]	72.2	8.0[c]	[71.3, 73.2]
60% (F)	70.2	7.2[a]	[69.3, 71.0]	58.7	9.6	[57.6, 59.8]	60.7	6.7[a]	[59.9, 61.5]	73.4	7.2[b]	[72.5, 74.2]	53.3	5.9	[52.6, 54.0]
70% (R)	60.7	7.7	[59.8, 61.6]	71.1	11.0[b]	[69.8, 72.4]	61.6	8.2[a]	[60.6, 62.5]	69.2	8.9[b]	[68.1, 70.2]	64.4	8.4[a]	[63.4, 65.4]
70% (T)	49.0	6.5[a]	[48.3, 49.8]	79.5	7.6[c]	[78.6, 80.4]	61.8	6.1[a]	[61.1, 62.5]	61.7	5.8[a]	[61.0, 62.4]	74.5	7.6[c]	[73.6, 75.4]
70% (F)	72.1	6.3[a]	[71.4, 72.9]	59.7	8.7	[58.6, 60.7]	61.9	5.3[a]	[61.3, 62.5]	77.0	6.3[c]	[76.3, 77.8]	53.6	4.7	[53.0, 54.1]

Table 12 Continued

Response insertion percentage	MLS			GIC			HPC			NUC			COG		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
80% (R)	61.5	8.1	[60.6, 62.5]	71.4	11.1[b]	[70.1, 72.7]	62.8	7.9[a]	[61.8, 63.7]	72.4	8.3[b]	[71.4, 73.4]	65.3	8.3[a]	[64.3, 66.3]
80% (T)	48.5	5.4[a]	[47.9, 49.2]	82.7	6.9[c]	[81.8, 83.5]	63.1	5.3[a]	[62.5, 63.4]	63.0	4.5[a]	[62.5, 63.6]	78.8	6.1[d]	[78.0, 79.5]
80% (F)	75.2	5.0[c]	[74.6, 75.8]	61.6	7.1[a]	[60.7, 62.4]	62.3	4.9[a]	[61.8, 62.9]	80.0	5.0[c]	[79.4, 80.6]	53.7	3.6	[53.3, 54.2]
90% (R)	62.0	8.1[a]	[61.0, 62.9]	73.7	9.8[b]	[72.6, 74.9]	63.0	8.4[a]	[62.0, 64.0]	73.7	7.5[b]	[72.8, 74.6]	68.0	8.2[b]	[67.0, 69.0]
90% (T)	47.6	3.9[a]	[47.1, 48.0]	85.5	4.5[d]	[85.0, 86.1]	64.4	3.4[b]	[64.0, 64.8]	63.6	3.3[a]	[63.2, 64.0]	81.8	4.3[e]	[81.2, 82.3]
90% (F)	77.5	3.7[c]	[77.1, 78.0]	62.4	5.6[a]	[61.8, 63.1]	64.0	3.3[a]	[63.6, 64.4]	83.0	3.8[c]	[82.6, 83.5]	54.2	2.9	[53.9, 54.6]
100% (R)	63.1	7.5[a]	[62.2, 64.0]	76.2	9.6[b]	[75.1, 77.4]	65.6	7.8[b]	[64.7, 66.6]	74.8	8.7[b]	[73.8, 75.8]	69.3	8.2[b]	[68.3, 70.3]
100% (T)	46.5	0.0[a]	[--, --]	88.3	0.0[d]	[--, --]	65.1	0.0[b]	[--, --]	64.6	0.0[a]	[--, --]	85.6	0.0[e]	[--, --]
100% (F)	79.9	0.0[c]	[--, --]	63.8	0.0[a]	[--, --]	65.1	0.0[b]	[--, --]	85.6	0.0[d]	[--, --]	54.4	0.0	[--, --]

Note. n = 277 and n = 275 for variable and acquiescent/counter-acquiescent response insertion, respectively. MLS = Malaise; GIC =

Table 12 Continued

Gastrointestinal Complaints; HPC = Head Pain Complaints; NUC = Neurological Complaints; COG = Cognitive Complaints; R = random; T = true; and F = false.

[a,b,c,d] The mean score differs from the 0% variable, acquiescent, or counter-acquiescent baseline mean score by one, two, three, or four times this scale's standard error of measurement, respectively.

When fixed acquiescent and counter-acquiescent responding reaches 100%, T-scores become constant and the standard deviations equal zero.

The format of results is as follows: Mean [lower, upper bounds of 95% Confidence Interval]. [--, --] = The confidence interval for this mean score could not be calculated due to the same reason described above. The inpatient sample (Archer, Griffin, & Aiduk, 1995) is the basis for all the results presented in this table.

Table 13

The Effects of Increasing Degrees of Variable, Fixed Acquiescent, and Fixed Counter-Acquiescent Response Insertion on Mean Inpatient Sample Internalizing (RCd-Associated) Scale T-Scores, Standard Deviations, and 95% Confidence Intervals.

Response insertion percentage	SUI			HLP			SFD			NFC		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
0% (R)	59.2	18.8	[56.9, 61.4]	50.3	12.1	[48.9, 51.7]	55.1	12.0	[53.7, 56.5]	51.8	10.6	[50.5, 53.0]
0% (T or F)	59.6	19.3	[57.3, 61.9]	50.5	12.4	[49.1, 52.0]	55.6	12.3	[54.1, 57.0]	52.1	10.9	[50.8, 53.4]
10% (R)	62.3	18.7	[60.1, 64.5]	51.8	11.6	[50.5, 53.2]	55.2	11.5	[53.8, 56.5]	52.6	9.8	[51.4, 53.8]
10% (T)	65.4	18.8	[63.2, 67.6]	53.3	12.2	[51.9, 54.8]	57.7	11.6	[56.3, 59.1]	54.7	10.4	[53.5, 56.0]
10% (F)	58.4	18.4	[56.3, 60.6]	50.5	11.5	[49.1, 51.8]	54.0	11.3	[52.6, 55.3]	50.8	10.0	[49.6, 51.9]
20% (R)	65.3	18.6	[63.2, 67.5]	52.6	11.1	[51.3, 54.0]	55.8	10.5	[54.5, 57.0]	53.0	9.0	[51.9, 54.1]
20% (T)	73.0	18.5[a]	[70.8, 75.2]	56.7	12.1	[55.2, 58.1]	59.6	10.6	[58.4, 60.9]	57.5	9.9	[56.3, 58.6]
20% (F)	57.0	16.9	[55.0, 59.0]	50.6	11.1	[49.3, 51.9]	52.5	10.6	[51.2, 53.7]	48.9	9.0	[47.8, 50.0]
30% (R)	66.8	17.4	[64.7, 68.8]	55.1	11.4	[53.8, 56.5]	55.8	9.6	[54.6, 56.9]	53.0	8.7	[52.0, 54.1]

Table 13 Continued

Response insertion percentage	SUI			HLP			SFD			NFC		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
30% (T)	78.3	17.9[a]	[76.2, 80.4]	59.9	11.7[a]	[58.5, 61.3]	60.8	10.4	[59.6, 62.1]	59.7	8.9[a]	[58.6, 60.7]
30% (F)	55.3	15.4	[53.4, 57.1]	50.6	10.1	[49.4, 51.8]	51.1	9.3	[50.0, 52.2]	47.8	8.2	[46.9, 48.8]
40% (R)	71.6	16.2[a]	[69.6, 73.5]	56.1	11.1	[54.8, 57.4]	56.1	9.7	[54.9, 57.2]	54.1	7.7	[53.2, 55.1]
40% (T)	84.8	15.5[b]	[83.0, 86.7]	61.5	11.5[a]	[60.2, 62.9]	62.6	10.2[a]	[61.4, 63.8]	61.9	9.0[a]	[60.8, 63.0]
40% (F)	54.4	14.8	[52.6, 56.2]	50.3	8.6	[49.2, 51.3]	49.8	8.2	[48.8, 50.7]	46.4	7.2[a]	[45.5, 47.2]
50% (R)	72.8	16.1[a]	[70.9, 74.7]	57.2	10.7	[55.9, 58.4]	56.5	9.7	[55.3, 57.6]	54.3	7.4	[53.4, 55.2]
50% (T)	91.1	15.1[c]	[89.3, 92.9]	66.2	11.0[b]	[64.9, 67.5]	65.2	8.9[a]	[64.1, 66.2]	64.9	8.7[b]	[63.9, 66.0]
50% (F)	53.6	13.3	[52.0, 55.2]	50.3	7.8	[49.3, 51.2]	48.6	7.7[a]	[47.7, 49.5]	44.9	6.7[a]	[44.2, 45.7]
60% (R)	75.0	15.0[a]	[73.2, 76.8]	58.9	10.7[a]	[57.7, 60.2]	56.9	8.5	[55.9, 57.9]	54.5	7.0	[53.7, 55.3]
60% (T)	95.9	14.0[c]	[94.3, 97.6]	67.0	9.8[b]	[65.8, 68.1]	66.5	8.8[a]	[65.4, 67.5]	68.3	7.7[b]	[67.4, 69.2]
60% (F)	51.4	11.1	[50.1, 52.8]	51.0	7.5	[50.1, 51.9]	48.2	7.6[a]	[47.3, 49.1]	43.2	6.4[a]	[42.4, 43.9]
70% (R)	79.0	15.9[a]	[77.1, 80.9]	60.7	10.1[a]	[59.5, 61.9]	56.9	8.3	[55.9, 57.9]	55.9	6.6	[55.2, 56.7]
70% (T)	101.6	13.0[d]	[100.1, 103.2]	70.1	9.1[b]	[69.0, 71.1]	68.4	8.2[b]	[67.5, 69.4]	71.3	6.9[c]	[70.5, 72.1]

Table 13 Continued

Response insertion percentage	SUI			HLP			SFD			NFC		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
70% (F)	49.9	9.9	[48.8, 51.1]	51.1	6.4	[50.3, 51.8]	46.1	6.2[a]	[45.4, 46.8]	42.2	5.8[a]	[41.6, 42.9]
80% (R)	81.2	15.7[b]	[79.3, 83.0]	61.8	9.5[a]	[60.7, 62.9]	57.2	8.0	[56.2, 58.1]	55.7	7.0	[54.9, 56.5]
80% (T)	106.2	10.3[d]	[105.0, 107.5]	72.9	7.6[c]	[72.0, 73.8]	71.3	6.5[b]	[70.5, 72.0]	74.2	5.8[d]	[73.5, 74.9]
80% (F)	48.3	8.0[a]	[47.3, 49.3]	51.1	5.8	[50.4, 51.8]	44.6	4.7[a]	[44.1, 45.2]	40.4	5.2[a]	[39.8, 41.0]
90% (R)	82.6	14.7[b]	[80.9, 84.3]	61.8	10.5[a]	[60.5, 63.0]	57.5	7.6	[56.6, 58.4]	56.9	6.4	[56.1, 57.6]
90% (T)	111.3	8.0[e]	[110.4, 112.3]	75.9	5.2[c]	[75.3, 76.5]	73.5	4.8[c]	[72.9, 74.1]	76.9	4.5[d]	[76.4, 77.5]
90% (F)	47.3	6.2[a]	[46.6, 48.1]	51.6	2.9	[51.2, 51.9]	43.7	4.2[b]	[43.2, 44.2]	38.0	3.4[a]	[37.6, 38.4]
100% (R)	83.6	16.2[b]	[81.7, 85.5]	64.3	10.9[a]	[63.0, 65.6]	57.4	7.5	[56.5, 58.2]	57.4	6.6[a]	[56.6, 58.2]
100% (T)	116.5	0.0[e]	[--, --]	78.5	0.0[c]	[--, --]	75.7	0.0[c]	[--, --]	80.4	0.0[e]	[--, --]
100% (F)	45.4	0.0[a]	[--, --]	51.6	0.0	[--, --]	41.8	0.0[b]	[--, --]	36.1	0.0[a]	[--, --]

Note. n = 277 and n = 275 for variable and acquiescent/counter-acquiescent response insertion, respectively.

Table 13 Continued

RCd = Demoralization; SUI = Suicidal/Death Ideation; HLP =Helplessness/Hopelessness; SFD = Self-Doubt; NFC = Inefficacy; R = random; T = true; and F = false.

[a,b,c,d,e] The mean score differs from the 0% variable, acquiescent, or counter-acquiescent baseline mean score by one, two, three, four, or five times this scale's standard error of measurement, respectively.

When fixed acquiescent and counter-acquiescent responding reaches 100%, T-scores become constant and the standard deviations equal zero.

The format of results is as follows: Mean [lower, upper bounds of 95% Confidence Interval]. [--, --] = The confidence interval for this mean score could not be calculated due to the same reason described above. The inpatient sample (Archer, Griffin, & Aiduk, 1995) is the basis for all the results presented in this table.

Table 14

The Effects of Increasing Degrees of Variable, Fixed Acquiescent, and Fixed Counter-Acquiescent Response Insertion on Mean Inpatient Sample Internalizing (RC7-Associated) Scale T-Scores, Standard Deviations, and 95% Confidence Intervals.

Response insertion	STW			AXY			ANP			BRF			MSF		
percentage	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
0% (R)	52.4	10.6	[51.2, 53.7]	53.4	12.3	[51.9, 54.8]	51.9	11.2	[50.5, 53.2]	52.0	10.6	[50.7, 53.3]	50.2	10.1	[49.0, 51.4]
0% (T or F)	52.6	10.7	[51.4, 53.9]	53.5	12.3	[52.0, 54.9]	51.9	11.2	[50.5, 53.2]	52.0	10.7	[50.7, 53.2]	50.3	10.3	[49.1, 51.5]
10% (R)	52.7	10.0	[51.5, 53.9]	55.6	12.8	[54.1, 57.2]	52.5	10.2	[51.3, 53.7]	55.1	10.5	[53.8, 56.3]	50.8	9.6	[49.7, 51.9]
10% (T)	53.9	10.3	[52.6, 55.1]	58.9	12.9	[57.3, 60.4]	53.1	10.6	[51.9, 54.4]	59.0	11.6	[57.6, 60.4]	50.5	9.4	[49.4, 51.6]
10% (F)	52.2	10.5	[51.0, 53.5]	52.7	11.7	[51.4, 54.1]	51.4	10.2	[50.2, 52.7]	52.3	10.7	[51.0, 53.6]	51.2	9.6	[50.0, 52.3]
20% (R)	53.4	9.5	[52.3, 54.5]	58.4	12.0	[56.9, 59.8]	52.9	9.7	[51.7, 54.0]	58.8	11.4	[57.5, 60.1]	50.9	8.9	[49.9, 52.0]
20% (T)	54.7	10.3	[53.5, 55.9]	64.6	12.9[a]	[63.0, 66.1]	54.7	9.6	[53.5, 55.8]	65.6	11.0[a]	[64.3, 66.9]	50.7	8.6	[49.6, 51.7]
20% (F)	51.7	9.6	[50.6, 52.8]	51.9	11.7	[50.5, 53.2]	51.6	9.4	[50.5, 52.7]	52.5	10.1	[51.3, 53.7]	51.4	8.6	[50.4, 52.4]
30% (R)	53.1	9.3	[51.9, 54.2]	61.7	12.6	[60.2, 63.1]	53.5	8.5	[52.5, 54.5]	61.9	10.7[a]	[60.7, 63.2]	51.4	8.1	[50.5, 52.4]
30% (T)	56.4	10.0	[55.2, 57.6]	69.4	13.2[a]	[67.9, 71.0]	56.3	8.7	[55.2, 57.3]	70.4	10.6[b]	[69.2, 71.7]	50.3	7.8	[49.4, 51.3]

Table 14 Continued

Response insertion percentage	STW *M*	STW *SD*	STW *95% CI*	AXY *M*	AXY *SD*	AXY *95% CI*	ANP *M*	ANP *SD*	ANP *95% CI*	BRF *M*	BRF *SD*	BRF *95% CI*	MSF *M*	MSF *SD*	MSF *95% CI*
30% (F)	50.7	8.9	[49.7, 51.8]	50.3	9.6	[49.2, 51.4]	51.6	8.2	[50.7, 52.6]	52.5	9.4	[51.4, 53.7]	52.2	7.6	[51.3, 53.1]
40% (R)	53.6	9.2	[52.5, 54.6]	62.5	13.3[a]	[60.9, 64.1]	54.3	8.5	[53.3, 55.3]	65.4	11.3[a]	[64.1, 66.8]	51.1	6.5	[50.4, 51.9]
40% (T)	57.0	9.0	[56.0, 58.1]	73.4	13.1[b]	[71.8, 74.9]	57.5	8.9	[56.4, 58.5]	77.0	11.7[c]	[75.6, 78.4]	51.0	6.6	[50.3, 51.8]
40% (F)	50.4	7.7	[49.5, 51.3]	50.3	9.7	[49.1, 51.5]	51.1	7.6	[50.2, 52.0]	53.0	9.2	[51.9, 54.0]	52.4	7.1	[51.6, 53.3]
50% (R)	53.1	8.7	[52.1, 54.2]	64.8	12.6[a]	[63.3, 66.3]	54.7	8.0	[53.8, 55.7]	67.7	10.5[a]	[66.5, 69.0]	51.8	6.5	[51.1, 52.6]
50% (T)	58.9	8.9	[57.9, 60.0]	79.2	12.0[b]	[77.7, 80.6]	58.7	8.0[a]	[57.8, 59.7]	80.9	10.9[c]	[79.6, 82.2]	50.9	5.7	[50.2, 51.5]
50% (F)	50.0	7.3	[49.2, 50.9]	48.8	8.4	[47.8, 49.8]	51.2	6.5	[50.4, 52.0]	53.8	8.9	[52.8, 54.9]	52.4	6.1	[51.7, 53.1]
60% (R)	54.2	8.5	[53.1, 55.2]	66.6	12.3[a]	[65.1, 68.1]	55.4	7.2	[54.6, 56.3]	69.8	12.1[b]	[68.4, 71.3]	51.5	6.4	[50.7, 52.3]
60% (T)	59.6	8.0	[58.7, 60.6]	81.8	11.8[c]	[80.4, 83.2]	59.6	6.9[a]	[58.8, 60.4]	86.3	11.4[d]	[85.0, 87.7]	50.8	5.3	[50.2, 51.4]
60% (F)	49.0	5.7	[48.4, 49.7]	48.3	7.9	[47.4, 49.3]	51.1	5.2	[50.4, 51.7]	54.2	7.5	[53.3, 55.1]	53.2	5.3	[52.5, 53.8]
70% (R)	55.4	8.8	[54.4, 56.4]	69.4	12.6[a]	[67.9, 70.9]	55.5	6.7	[54.7, 56.3]	73.4	11.3[b]	[72.1, 74.7]	51.8	5.8	[51.1, 52.5]
70% (T)	60.7	7.0[a]	[59.8, 61.5]	88.3	10.4[d]	[87.1, 89.6]	60.7	5.9[a]	[60.0, 61.4]	91.8	9.6[e]	[90.7, 92.9]	50.8	4.1	[50.3, 51.3]
70% (F)	48.8	5.0	[48.2, 49.4]	47.4	7.2	[46.5, 48.3]	51.1	4.3	[50.6, 51.6]	54.3	7.3	[53.5, 55.2]	53.7	4.5	[53.2, 54.2]

Table 14 Continued

Response insertion percentage	STW			AXY			ANP			BRF			MSF		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
80% (R)	54.5	8.4	[53.5, 55.5]	69.9	13.0[a]	[68.3, 71.4]	56.2	7.4	[55.4, 57.1]	77.0	11.1[c]	[75.7, 78.3]	52.6	5.9	[51.9, 53.3]
80% (T)	62.2	6.7[a]	[61.4, 63.0]	92.9	8.8[d]	[91.9, 94.0]	62.0	5.9[a]	[61.3, 62.7]	98.4	8.5[e]	[97.4, 99.4]	50.9	3.0	[50.6, 51.3]
80% (F)	48.2	4.5	[47.6, 48.7]	46.1	5.3	[45.4, 46.7]	50.9	4.1	[50.4, 51.4]	54.6	6.1	[53.8, 55.3]	53.2	3.4	[52.8, 53.6]
90% (R)	55.3	8.0	[54.4, 56.3]	74.5	11.6[b]	[73.2, 75.9]	56.3	6.6	[55.5, 57.1]	79.8	12.2[c]	[78.3, 81.2]	53.2	5.6	[52.5, 53.8]
90% (T)	63.9	4.2[a]	[63.4, 64.4]	96.6	6.5[e]	[95.9, 97.4]	64.0	4.5[b]	[63.4, 64.5]	103.0	6.2[f]	[102.3, 103.7]	50.9	2.1	[50.7, 51.2]
90% (F)	47.5	2.8	[47.1, 47.8]	45.0	4.3	[44.5, 45.5]	51.3	2.1	[51.0, 51.5]	55.4	4.0	[54.9, 55.8]	53.7	2.1	[53.5, 54.0
100% (R)	56.4	8.0	[55.4, 57.3]	75.8	11.8[b]	[74.4, 77.2]	57.6	6.9[a]	[56.7, 58.4]	81.6	11.1[c]	[80.3, 83.0]	53.3	5.8	[52.6, 54.0]
100% (T)	65.2	0.0[a]	[--, --]	101.5	0.0[e]	[--, --]	65.8	0.0[b]	[--, --]	108.9[h]	0.0	[--, --]	51.0	0.0	[--, --]
100% (F)	47.4	0.0	[--, --]	44.0	0.0[a]	[--, --]	51.0	0.0	[--, --]	55.8	0.0	[--, --]	54.1	0.0	[--, --]

Note. n = 277 and n = 275 for variable and acquiescent/counter-acquiescent response insertion, respectively. RC7 = Dysfunctional

Table 14 Continued

Negative Emotions; STW = Stress/Worry; AXY = Anxiety; ANP = Anger-Proneness; BRF = Behavior-Restricting Fears; MSF = Multiple Specific Fears; R = random; T = true; and F = false.

[a,b,c,d,e,f,g,h] The mean score differs from the 0% variable, acquiescent, or counter-acquiescent baseline mean score by one, two, three, four, five, six, seven, or eight times this scale's standard error of measurement, respectively.

When fixed acquiescent and counter-acquiescent responding reaches 100%, T-scores become constant and the standard deviations equal zero.

The format of results is as follows: Mean [lower, upper bounds of 95% Confidence Interval]. [--, --] = The confidence interval for this mean score could not be calculated due to the same reason described above. The inpatient sample (Archer, Griffin, & Aiduk, 1995) is the basis for all the results presented in this table.

Table 15

The Effects of Increasing Degrees of Variable, Fixed Acquiescent, and Fixed Counter-Acquiescent Response Insertion on Mean Inpatient Sample Externalizing Scale T-Scores, Standard Deviations, and 95% Confidence Intervals.

Response insertion	JCP			SUB			AGG			ACT		
percentage	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
0% (R)	58.2	13.3	[56.6, 59.8]	56.5	14.9	[54.8, 58.3]	51.2	11.3	[49.8, 52.5]	48.4	10.9	[47.1, 49.7]
0% (T or F)	58.2	13.3	[56.6, 59.8]	56.3	14.9	[54.5, 58.1]	51.2	11.3	[49.8, 52.5]	48.5	10.9	[47.2, 49.8]
10% (R)	58.6	12.4	[57.1, 60.1]	57.5	13.8	[55.9, 59.2]	52.5	11.0	[51.2, 53.8]	49.2	10.0	[48.0, 50.3]
10% (T)	61.0	12.3	[59.5, 62.4]	58.9	13.9	[57.2, 60.5]	55.6	11.3	[54.3, 57.0]	51.3	10.9	[50.0, 52.6]
10% (F)	56.4	12.4	[55.0, 57.9]	55.3	13.8	[53.6, 56.9]	49.8	10.5	[48.6, 51.1]	46.9	9.8	[45.8, 48.1]
20% (R)	58.8	11.5	[57.4, 60.2]	58.6	13.5	[57.0, 60.1]	53.9	10.6	[52.7, 55.2]	49.5	10.3	[48.3, 50.7]
20% (T)	64.1	11.3	[62.7, 65.4]	61.5	13.3	[59.9, 63.1]	59.6	10.5[a]	[58.3, 60.8]	54.4	11.1	[53.1, 55.7]
20% (F)	54.7	12.0	[53.3, 56.2]	55.2	12.9	[53.6, 56.7]	49.1	9.8	[47.9, 50.2]	45.5	9.3	[44.4, 46.6]
30% (R)	60.2	11.2	[58.9, 61.5]	60.3	13.0	[58.7, 61.8]	55.0	9.7	[53.8, 56.1]	49.8	9.2	[48.7, 50.9]

Table 15 Continued

Response insertion percentage	JCP *M*	JCP *SD*	JCP *95% CI*	SUB *M*	SUB *SD*	SUB *95% CI*	AGG *M*	AGG *SD*	AGG *95% CI*	ACT *M*	ACT *SD*	ACT *95% CI*
30% (T)	66.4	11.2[a]	[65.1, 67.8]	63.9	12.2[a]	[62.4, 65.3]	63.1	10.6[a]	[61.9, 64.4]	57.6	10.7[a]	[56.3, 58.9]
30% (F)	53.1	11.1	[51.8, 54.4]	54.1	11.8	[52.7, 55.5]	47.2	9.0	[46.2, 48.3]	43.8	8.4	[42.8, 44.8]
40% (R)	60.4	10.3	[59.1, 61.6]	60.3	11.1	[59.0, 61.6]	56.8	9.5	[55.7, 57.9]	50.0	9.5	[48.9, 51.1]
40% (T)	68.2	10.6[a]	[67.0, 69.5]	67.2	11.9[a]	[65.8, 68.6]	66.8	9.4[b]	[65.7, 67.9]	60.7	11.0[a]	[59.4, 62.0]
40% (F)	51.9	9.7	[50.8, 53.1]	53.4	9.7	[52.2, 54.5]	45.9	8.0	[44.9, 46.8]	41.7	6.6[a]	[40.9, 42.5]
50% (R)	60.5	9.8	[59.3, 61.6]	60.4	10.8	[59.1, 61.7]	57.9	8.8[a]	[56.9, 58.9]	50.7	8.6	[49.7, 51.7]
50% (T)	71.1	9.2[a]	[70.0, 72.2]	68.9	11.6[a]	[67.5, 70.3]	70.4	9.7[c]	[69.2, 71.5]	64.6	11.1[b]	[63.3, 65.9]
50% (F)	49.9	9.1[a]	[48.8, 50.9]	52.6	8.7	[51.6, 53.7]	44.5	7.6[a]	[43.6, 45.4]	40.6	6.0[a]	[39.9, 41.3]
60% (R)	61.6	8.7	[60.6, 62.7]	61.7	10.3	[60.5, 62.9]	58.9	9.2[a]	[57.8, 59.9]	51.8	8.2	[50.8, 52.7]
60% (T)	73.7	8.0[b]	[72.8, 74.7]	72.6	10.5[b]	[71.4, 73.9]	75.2	8.9[c]	[74.2, 76.3]	68.8	10.4[c]	[67.5, 70.0]
60% (F)	47.8	7.8[a]	[46.8, 48.7]	52.0	7.7	[51.0, 52.9]	43.3	6.6[a]	[42.6, 44.1]	39.3	6.0[a]	[38.6, 40.0]
70% (R)	61.7	9.0	[60.7, 62.8]	61.6	9.6	[60.4, 62.7]	60.4	8.9[a]	[59.4, 61.5]	52.8	8.3	[51.8, 53.7]
70% (T)	75.9	7.0[b]	[75.1, 76.8]	75.4	9.7[b]	[74.3, 76.6]	79.2	7.8[d]	[78.3, 80.2]	71.5	9.1[c]	[70.4, 72.5]

Table 15 Continued

Response insertion percentage	JCP *M*	JCP *SD*	JCP *95% CI*	SUB *M*	SUB *SD*	SUB *95% CI*	AGG *M*	AGG *SD*	AGG *95% CI*	ACT *M*	ACT *SD*	ACT *95% CI*
70% (F)	45.6	7.2[a]	[44.8, 46.5]	51.7	5.9	[51.0, 52.4]	41.9	6.3[a]	[41.1, 42.6]	37.9	5.0[a]	[37.3, 38.5]
80% (R)	61.3	8.5	[60.3, 62.3]	64.4	10.0[a]	[63.3, 65.5]	61.7	8.7[a]	[60.7, 62.7]	52.2	7.9	[51.3, 53.1]
80% (T)	78.2	5.9[b]	[77.5, 78.9]	78.5	7.5[c]	[77.6, 79.4]	82.9	6.7[e]	[82.1, 83.7]	75.1	7.5[d]	[74.2, 76.0]
80% (F)	44.8	6.5[a]	[44.0, 45.5]	50.9	5.2	[50.3, 51.5]	40.4	4.9[a]	[39.8, 41.0]	36.7	4.2[a]	[36.2, 37.2]
90% (R)	62.4	8.3	[61.4, 63.4]	63.5	9.6[a]	[62.4, 64.7]	64.0	8.5[b]	[63.0, 65.0]	53.9	8.4	[52.9, 54.9]
90% (T)	80.8	4.1[c]	[80.3, 81.3]	80.9	5.8[c]	[80.2, 81.6]	87.4	4.6[e]	[86.8, 87.9]	79.2	5.7[d]	[78.5, 79.8]
90% (F)	42.3	4.7[b]	[41.7, 42.8]	50.7	2.8	[50.3, 51.0]	39.1	3.7[a]	[38.6, 39.5]	35.0	3.1[b]	[34.6, 35.4]
100% (R)	63.5	8.9	[62.4, 64.5]	65.3	9.9[a]	[64.1, 66.5]	64.9	9.0[b]	[63.9, 66.0]	53.4	8.5	[52.4, 54.4]
100% (T)	83.7	0.0[c]	[--, --]	84.8	0.0[d]	[--, --]	91.5	0.0[f]	[--, --]	83.5	0.0[e]	[--, --]
100% (F)	40.1	0.0[b]	[--, --]	50.3	0.0	[--, --]	37.3	0.0[b]	[--, --]	33.2	0.0[b]	[--, --]

Note. n = 277 and n = 275 for variable and acquiescent/counter-acquiescent response insertion, respectively.

Table 15 Continued

JCP = Juvenile Conduct Problems; SUB = Substance Abuse; AGG = Aggression; ACT = Activation; R = random; T = true; and F = false.

[a,b,c,d,e,f] The mean score differs from the 0% variable, acquiescent, or counter-acquiescent baseline mean score by one, two, three, four, five, or six times this scale's standard error of measurement, respectively. When fixed acquiescent and counter-acquiescent responding reaches 100%, T-scores become constant and the standard deviations equal zero.

The format of results is as follows: Mean [lower, upper bounds of 95% Confidence Interval]. [--, --] = The confidence interval for this mean score could not be calculated due to the same reason described above. The inpatient sample (Archer, Griffin, & Aiduk, 1995) is the basis for all the results presented in this table.

Table 16

The Effects of Increasing Degrees of Variable, Fixed Acquiescent, and Fixed Counter-Acquiescent Response Insertion on Mean Inpatient Sample Interpersonal Scale T-Scores.

Response insertion	FML			IPP			SAV			SHY			DSF		
percentage	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
0% (R)	53.9	13.1	[52.3, 55.4]	48.8	10.3	[47.6, 50.0]	50.7	10.9	[49.4, 51.9]	50.4	10.5	[49.2, 51.7]	50.2	10.8	[49.0, 51.5]
0% (T or F)	54.1	13.4	[52.5, 55.7]	48.9	10.5	[47.7, 50.2]	50.6	11.0	[49.2, 51.9]	50.4	10.5	[49.1, 51.6]	49.7	10.3	[48.5, 51.0]
10% (R)	54.8	11.9	[53.4, 56.2]	49.5	9.7	[48.4, 50.6]	51.4	10.1	[50.2, 52.6]	50.3	9.5	[49.1, 51.4]	53.6	11.8	[52.2, 55.0]
10% (T)	56.0	12.4	[54.5, 57.5]	47.6	9.5	[46.5, 48.7]	49.4	9.5	[48.3, 50.5]	51.9	10.1	[50.7, 53.1]	55.4	12.0	[54.0, 56.8]
10% (F)	54.0	12.2	[52.5, 55.4]	51.0	10.1	[49.8, 52.2]	52.9	10.4	[51.7, 54.1]	49.5	9.5	[48.4, 50.7]	50.5	10.4	[49.3, 51.8]
20% (R)	56.0	11.3	[54.6, 57.3]	49.4	8.6	[48.3, 50.4]	52.0	8.8	[50.9, 53.0]	51.0	8.3	[50.0, 51.9]	56.9	11.9	[55.5, 58.3]
20% (T)	58.3	11.8	[56.9, 59.7]	46.5	8.1	[45.5, 47.4]	48.7	8.8	[47.6, 49.7]	53.1	9.4	[51.9, 54.2]	60.9	12.3[a]	[59.4, 62.3]
20% (F)	54.2	11.3	[52.9, 55.5]	53.0	9.6	[51.9, 54.2]	55.3	9.2	[54.2, 56.4]	48.6	7.9	[47.6, 49.5]	51.5	10.0	[50.3, 52.7]
30% (R)	56.4	10.2	[55.2, 57.6]	49.6	8.0	[48.7, 50.5]	52.7	8.4	[51.7, 53.7]	50.9	8.2	[49.9, 51.9]	59.8	12.4[a]	[58.3, 61.2]
30% (T)	60.3	10.9	[59.0, 61.6]	45.6	8.1	[44.6, 46.6]	47.8	8.0	[46.8, 48.7]	53.8	8.4	[52.8, 54.8]	65.8	12.4[b]	[64.3, 67.3]

Table 16 Continued

Response insertion percentage	FML *M*	FML *SD*	FML *95% CI*	IPP *M*	IPP *SD*	IPP *95% CI*	SAV *M*	SAV *SD*	SAV *95% CI*	SHY *M*	SHY *SD*	SHY *95% CI*	DSF *M*	DSF *SD*	DSF *95% CI*
30% (F)	54.0	11.0	[52.7, 55.3]	54.5	9.4[a]	[53.4, 55.7]	57.2	8.9[a]	[56.2, 58.3]	47.2	7.4	[46.3, 48.1]	52.1	10.1	[50.9, 53.3]
40% (R)	57.3	10.0	[56.1, 58.5]	50.4	7.8	[49.5, 51.3]	53.0	7.5	[52.1, 53.9]	51.0	7.3	[50.2, 51.9]	62.2	12.1[a]	[60.7, 63.6]
40% (T)	61.6	9.8[a]	[60.5, 62.8]	44.5	6.9	[43.7, 45.4]	47.3	7.3	[46.4, 48.2]	55.4	8.4[a]	[54.4, 56.4]	71.5	11.9[c]	[70.0, 72.9]
40% (F)	53.5	9.6	[52.4, 54.7]	57.4	9.1[a]	[56.3, 58.5]	58.8	8.2[a]	[57.8, 59.7]	46.8	6.2	[46.1, 47.6]	53.5	9.3	[52.3, 54.6]
50% (R)	59.0	9.0	[57.9, 60.0]	51.1	7.3	[50.3, 52.0]	53.7	6.6	[52.9, 54.4]	51.3	6.8	[50.5, 52.1]	64.9	11.9[b]	[63.5, 66.3]
50% (T)	63.9	9.0[a]	[62.8, 64.9]	43.5	6.2[a]	[42.7, 44.2]	46.9	6.6	[46.1, 47.6]	56.5	7.9[a]	[55.6, 57.5]	75.3	11.4[c]	[74.0, 76.7]
50% (F)	54.0	8.4	[53.0, 55.0]	59.2	8.9[a]	[58.1, 60.2]	61.5	7.8[b]	[60.5, 62.4]	45.8	5.7	[45.2, 46.5]	54.4	9.1	[53.3, 55.5]
60% (R)	59.1	9.1	[58.0, 60.2]	51.2	7.1	[50.4, 52.1]	54.3	6.9	[53.5, 55.1]	51.2	5.8	[50.5, 51.9]	68.1	13.2[b]	[66.5, 69.6]
60% (T)	66.1	8.3[a]	[65.1, 67.0]	42.7	4.9[a]	[42.2, 43.3]	46.0	5.8	[45.3, 46.7]	58.3	7.7[a]	[57.4, 59.2]	79.5	11.4[d]	[78.2, 80.9]
60% (F)	53.8	7.2	[53.0, 54.7]	62.1	8.3[b]	[61.1, 63.1]	63.5	7.4[b]	[62.6, 64.3]	45.4	4.7[a]	[44.8, 46.0]	54.7	7.9	[53.8, 55.7]
70% (R)	61.1	8.3[a]	[60.2, 62.1]	52.1	7.1	[51.3, 53.0]	55.0	6.1	[54.2, 55.7]	51.4	5.3	[50.7, 52.0]	70.6	12.5[b]	[69.2, 72.1]
70% (T)	67.4	7.1[b]	[66.6, 68.2]	42.1	4.4[a]	[41.6, 42.6]	44.9	4.9[a]	[44.3, 45.5]	59.0	6.6[a]	[58.2, 59.8]	85.1	9.1[e]	[84.0, 86.2]
70% (F)	53.7	5.7	[53.0, 54.3]	65.2	7.3[c]	[64.3, 66.1]	66.3	7.3[c]	[65.4, 67.2]	45.0	3.8[a]	[44.6, 45.5]	55.4	7.6	[54.5, 56.3]

Table 16 Continued

Response insertion percentage	FML			IPP			SAV			SHY			DSF		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
80% (R)	60.2	7.5	[59.3, 61.1]	52.2	6.4	[51.5, 53.0]	54.8	6.0	[54.1, 55.5]	51.9	6.0	[51.2, 52.6]	72.2	12.5[c]	[70.7, 73.6]
80% (T)	69.4	5.7[b]	[68.7, 70.0]	41.0	3.8[a]	[40.5, 41.4]	44.6	3.7[a]	[44.2, 45.1]	61.6	5.8[b]	[60.9, 62.3]	88.8	8.8[e]	[87.7, 89.8]
80% (F)	53.9	4.8	[53.3, 54.5]	67.1	6.3[c]	[66.3, 67.8]	68.9	5.9[c]	[68.2, 69.6]	44.8	3.2[a]	[44.4, 45.1]	56.2	6.7	[55.4, 57.0]
90% (R)	61.8	8.0[a]	[60.9, 62.8]	52.4	6.4	[51.6, 53.2]	55.9	5.7[a]	[55.3, 56.6]	51.5	5.5	[50.9, 52.2]	75.8	11.6[c]	[74.5, 77.2]
90% (T)	71.6	4.5[b]	[71.0, 72.1]	40.1	2.9[a]	[39.7, 40.4]	43.7	2.7[a]	[43.4, 44.0]	63.1	4.7[b]	[62.6, 63.7]	92.6	7.3[f]	[91.8, 93.5]
90% (F)	53.5	3.1	[53.2, 53.9]	71.0	4.9[d]	[70.4, 71.5]	71.9	4.4[d]	[71.4, 72.4]	44.2	2.8[a]	[43.9, 44.5]	57.4	4.9[a]	[56.9, 58.0]
100% (R)	62.8	7.6[a]	[61.9, 63.7]	52.8	6.2	[52.0, 53.5]	56.5	5.7[a]	[55.9, 57.2]	51.3	4.9	[50.7, 51.9]	77.5	12.2[c]	[76.1, 79.0]
100% (T)	73.6	0.0[c]	[--, --]	39.1	0.0[a]	[--, --]	42.9	0.0[a]	[--, --]	65.6	0.0[c]	[--, --]	97.8	0.0[f]	[--, --]
100% (F)	53.3	0.0	[--, --]	74.4	0.0[d]	[--, --]	75.2	0.0[e]	[--, --]	43.8	0.0[a]	[--, --]	58.3	0.0[a]	[--, --]

Note. n = 277 and n = 275 for variable and acquiescent/counter-acquiescent response insertion, respectively. FML = Family

Table 16 Continued

Problems; IPP = Interpersonal Passivity; SAV = Social Avoidance; SHY = Shyness; DSF = Disaffiliativeness; R = random; T = true; F = false.

[a,b,c,d,e,f] The mean score differs from the 0% variable, acquiescent, or counter-acquiescent baseline mean score by one, two, three, four, five, or six times this scale's standard error of measurement, respectively.

When fixed acquiescent and counter-acquiescent responding reaches 100%, T-scores become constant and the standard deviations equal zero.

The format of results is as follows: Mean [lower, upper bounds of 95% Confidence Interval]. [--, --] = The confidence interval for this mean score could not be calculated due to the same reason described above. The inpatient sample (Archer, Griffin, & Aiduk, 1995) is the basis for all the results presented in this table.

Table 17

The Effects of Increasing Degrees of Variable, Fixed Acquiescent, and Fixed Counter-Acquiescent Response Insertion on Mean PSY-5 Inpatient Sample Scale T-Scores.

Response insertion	AGGR-r			PSYC-r			DISC-r			NEGE-r			INTR-r		
percentage	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*	*M*	*SD*	*95% CI*
0% (R)	51.2	10.1	[50.0, 52.4]	50.7	10.4	[49.5, 51.9]	54.5	11.9	[53.1, 55.9]	53.8	10.8	[52.5, 55.1]	52.2	11.3	[50.8, 53.5]
0% (T or F)	51.1	10.2	[49.9, 52.3]	50.6	10.4	[49.3, 51.8]	54.5	12.0	[53.1, 55.9]	54.0	10.9	[52.7, 55.3]	52.1	11.4	[50.7, 53.4]
10% (R)	51.0	9.5	[49.9, 52.1]	55.7	9.3	[54.6, 56.8]	54.8	10.8	[53.5, 56.1]	54.4	10.1	[53.2, 55.6]	53.1	10.5	[51.9, 54.4]
10% (T)	53.5	10.2	[52.3, 54.7]	60.9	8.8[a]	[59.8, 61.9]	57.1	11.3	[55.7, 58.4]	55.7	10.3	[54.5, 56.9]	49.9	10.0	[48.7, 51.1]
10% (F)	48.9	9.2	[47.8, 49.9]	50.2	9.7	[49.0, 51.3]	53.0	10.7	[51.8, 54.3]	53.2	10.2	[52.0, 54.4]	55.6	11.1	[54.3, 56.9]
20% (R)	51.0	8.7	[49.9, 52.0]	60.9	7.9[a]	[59.9, 61.8]	55.4	10.0	[54.2, 56.5]	55.1	9.2	[54.0, 56.1]	54.1	9.9	[53.0, 55.3]
20% (T)	55.6	9.7	[54.4, 56.7]	68.9	8.7[c]	[67.9, 69.9]	59.6	10.3	[58.3, 60.8]	57.9	10.0	[56.7, 59.1]	48.0	9.1	[46.9, 49.1]
20% (F)	47.0	7.7	[46.0, 47.9]	49.7	9.1	[48.6, 50.8]	51.7	9.8	[50.5, 52.8]	52.6	8.7	[51.5, 53.6]	59.3	10.5[a]	[58.1, 60.6]
30% (R)	50.9	8.3	[49.9, 51.9]	64.0	8.9[b]	[62.9, 65.0]	56.1	9.3	[55.0, 57.2]	55.4	8.9	[54.3, 56.4]	54.7	8.8	[53.6, 55.7]
30% (T)	57.8	10.2[a]	[56.6, 59.1]	75.4	8.7[d]	[74.3, 76.4]	62.2	9.9[a]	[61.0, 63.4]	59.8	9.7[a]	[58.7, 61.0]	46.1	8.1[a]	[45.2, 47.1]

Table 17 Continued

Response insertion percentage	AGGR-r *M*	AGGR-r *SD*	AGGR-r *95% CI*	PSYC-r *M*	PSYC-r *SD*	PSYC-r *95% CI*	DISC-r *M*	DISC-r *SD*	DISC-r *95% CI*	NEGE-r *M*	NEGE-r *SD*	NEGE-r *95% CI*	INTR-r *M*	INTR-r *SD*	INTR-r *95% CI*
30% (F)	45.4	7.0[a]	[44.6, 46.3]	49.4	8.4	[48.4, 50.4]	50.0	8.5	[49.0, 51.0]	51.5	7.9	[50.5, 52.4]	63.7	10.3[b]	[62.5, 65.0]
40% (R)	50.9	8.2	[50.0, 51.0]	67.2	8.4[b]	[66.2, 68.1]	57.1	8.3	[56.2, 58.1]	56.0	8.2	[55.0, 57.0]	55.3	8.4	[54.3, 56.3]
40% (T)	59.8	9.2[a]	[58.7, 60.9]	85.4	8.7[e]	[84.4, 86.4]	64.4	9.9[a]	[63.2, 65.6]	61.9	9.4[a]	[60.8, 63.0]	44.2	7.4[a]	[43.3, 45.1]
40% (F)	43.2	5.5[a]	[42.5, 43.8]	48.6	7.8	[47.7, 49.6]	48.9	7.9[a]	[48.0, 49.9]	50.8	7.1	[50.0, 51.7]	67.6	9.0[b]	[66.5, 68.6]
50% (R)	50.2	6.9	[49.4, 51.1]	71.0	7.9[c]	[70.1, 72.0]	57.0	8.5	[56.0, 58.0]	56.2	7.9	[55.3, 57.1]	56.4	7.6	[55.5, 57.3]
50% (T)	62.4	9.2[b]	[61.3, 63.5]	91.8	9.0[f]	[90.7, 92.9]	67.4	8.4[b]	[66.4, 68.4]	64.2	8.6[a]	[63.2, 65.3]	42.9	6.8[a]	[42.1, 43.7]
50% (F)	42.1	5.0[a]	[41.5, 42.7]	48.8	7.1	[48.0, 49.7]	47.6	6.5[a]	[46.8, 48.4]	50.0	5.9	[49.3, 50.7]	71.8	8.7[c]	[70.8, 72.8]
60% (R)	50.4	6.6	[49.7, 51.2]	75.6	8.3[d]	[74.6, 76.6]	58.2	7.3	[57.3, 59.1]	56.8	7.0	[56.0, 57.6]	57.1	8.0	[56.2, 58.0]
60% (T)	65.7	8.7[b]	[64.7, 66.8]	99.6	8.0[h]	[98.6, 100.5]	70.7	7.7[b]	[69.8, 71.6]	66.1	7.6[b]	[65.2, 67.0]	40.6	5.8[b]	[39.9, 41.2]
60% (F)	40.4	4.0[b]	[39.9, 40.9]	48.6	6.4	[47.9, 49.4]	46.3	5.3[a]	[45.7, 46.9]	49.0	4.6	[48.4, 49.5]	75.8	7.5[d]	[74.9, 76.7]
70% (R)	50.3	6.6	[49.5, 51.1]	77.8	8.0[d]	[76.9, 78.8]	58.3	7.1	[57.5, 59.1]	57.7	7.1	[56.8, 58.5]	57.8	7.0[a]	[56.9, 58.6]
70% (T)	68.0	7.7[c]	[67.1, 68.9]	106.5	7.5[i]	[105.6, 107.3]	73.5	6.2[c]	[72.6, 74.1]	68.7	6.4[b]	[68.0, 69.5]	38.9	4.9[b]	[38.3, 39.4]
70% (F)	39.1	3.9[b]	[38.6, 39.6]	48.0	5.5	[47.4, 48.6]	44.8	4.5[a]	[44.3, 45.3]	48.7	4.0[a]	[48.2, 49.1]	80.4	6.5[e]	[79.6, 81.2]

Table 17 Continued

Response insertion percentage	AGGR-r *M*	AGGR-r *SD*	AGGR-r *95% CI*	PSYC-r *M*	PSYC-r *SD*	PSYC-r *95% CI*	DISC-r *M*	DISC-r *SD*	DISC-r *95% CI*	NEGE-r *M*	NEGE-r *SD*	NEGE-r *95% CI*	INTR-r *M*	INTR-r *SD*	INTR-r *95% CI*
80% (R)	50.5	6.4	[49.7, 51.2]	81.9	8.5[e]	[80.9, 82.9]	58.0	6.7	[57.2, 58.8]	57.7	6.6	[56.9, 58.4]	58.3	7.0[a]	[57.5, 59.1]
80% (T)	71.6	6.5[d]	[70.9, 72.4]	114.1	6.2[i]	[113.4, 114.9]	76.0	5.0[c]	[75.4, 76.6]	71.5	5.8[c]	[70.8, 72.2]	36.9	4.1[b]	[36.4, 37.4]
80% (F)	37.9	3.0[b]	[37.6, 38.3]	47.8	4.7	[47.3, 48.4]	44.0	3.8[a]	[43.6, 44.5]	47.8	3.1[a]	[47.4, 48.1]	84.3	5.6[f]	[83.6, 84.9]
90% (R)	51.0	6.4	[50.2, 51.7]	85.5	8.6[e]	[84.5, 86.5]	59.3	6.2	[58.5, 60.0]	58.8	7.0	[58.0, 59.6]	59.8	7.0[a]	[59.0, 60.6]
90% (T)	75.1	4.6[d]	[74.5, 76.5]	112.4	4.7[i]	[121.8, 122.9]	78.5	4.0[d]	[78.0, 79.0]	74.1	4.1[c]	[73.6, 74.6]	34.7	3.2[c]	[34.3, 35.1]
90% (F)	36.3	2.4[b]	[36.0, 36.6]	47.2	3.0	[46.9, 47.6]	42.7	2.4[b]	[42.4, 43.0]	47.4	1.9[a]	[47.2, 47.6]	88.9	3.8[g]	[88.5, 89.4]
100% (R)	50.6	6.0	[49.9, 51.3]	89.1	7.7[f]	[88.2, 90.1]	59.7	6.7	[58.9, 60.5]	59.5	6.3[a]	[58.7, 60.2]	61.3	7.0[a]	[60.4, 62.1]
100% (T)	78.5	0.0[e]	[--, --]	129.7	0.0[i]	[--, --]	81.9	0.0[e]	[--, --]	76.7	0.0[d]	[--, --]	32.2	0.0[c]	[--, --]
100% (F)	34.6	0.0[c]	[--, --]	46.6	0.0	[--, --]	41.5	0.0[b]	[--, --]	46.9	0.0[a]	[--, --]	93.3	0.0[g]	[--, --]

Note. n = 277 and n = 275 for variable and acquiescent/counter-acquiescent response insertion, respectively PSY-5 = Personality Psychopathology Five; AGGR-r = Aggressiveness-Revised; PSYC-r = Psychoticism-Revised; DISC-r = Disconstraint-Revised;

Table 17 Continued

NEGE-r = Negative Emotionality/Neuroticism-Revised; INTR-r = Introversion/Low Positive Emotions-Revised; R = random; T = true; and F = false.

[a,b,c,d,e,f,g,h,i] The mean score differs from the 0% variable, acquiescent, or counter-acquiescent baseline mean score by one, two, three, four, five, six, seven, eight, and nine or more times this scale's standard error of measurement, respectively.

When fixed acquiescent and counter-acquiescent responding reaches 100%, T-scores become constant and the standard deviations equal zero.

The format of results is as follows: Mean [lower, upper bounds of 95% Confidence Interval]. [--, --] = The confidence interval for this mean score could not be calculated due to the same reason described above. The inpatient sample (Archer, Griffin, & Aiduk, 1995) is the basis for all the results presented in this table.

As with the normative sample analyses, these tables include SP and PSY-5 Scale mean T-scores for baseline (i.e., 0% insertion for random, acquiescent, and counter-acquiescent responding) and response insertion conditions (i.e., 10-100% insertion for random, acquiescent, and counter-acquiescent responding). Only the results for the random response insertion analyses will be discussed in this section; results of acquiescent and counter-acquiescent responding will be discussed subsequently. These tables also include mean score standard deviations; alphabetical superscripts to indicate the magnitude of the deviation, as indicated by multiples of SEMs, between a SP or PSY-5 Scale baseline mean T-score and a response insertion mean T-score; and 95% confidence intervals for each mean T-score. Further, should be noted that the Internalizing SP scales are divided into two separate tables: Table 8 presents the SP Scale T-scores associated with the construct of demoralization, as represented by the RCd Scale, while Table 9 presents the T-scores associated with the construct of dysfunctional negative emotions, as represented by the RC7 Scale.

For 16 of the 28 SP and PSY-5 Scales, increasing degrees of simulated random responding resulted in a monotonic increase in scale mean T-scores. For 11 of the SP and PSY-5 scales, which included NFC, STW, MSF, JCP, SUB, ACT, FML, IPP, SAV, SHY, and DISC-r, increasing degrees of random responding resulted in a generally monotonic increase in scale mean T-scores, although this increase was not consistent from baseline to 100% response insertion. For example, the mean T-score for STW at 50% (53.1T) is lower than the mean at 40% (53.6T). Similarly, increasing degrees of random responding resulted in a generally monotonic decrease in mean T-scores for AGGR-r, although this decrease was not consistent.

As with the normative sample analyses, SP and PSY-5 mean T-scores differed in their susceptibility to random response insertion. These differences are discussed in the context of the same three indicators described in the Procedures section and previously used for the normative sample results. For the random response insertion analyses, a change of ≥ 5 T-score points was observed for the following SP and PSY-5 Scales. Somatic/Cognitive Scale scores increased by ≥ 5 T-score points at random response insertion rates ranging from 20% (GIC and NUC) to 50% (COG). Internalizing Scale T-scores increased at rates ranging from 20% (SUI, AXY, BRF) and, for the SFD, STW, and MSF Scales, did not change by ≥ 5 T-score points at 100% random response insertion. Externalizing Scale T-scores increased at rates ranging from 40% (AGG) to 100% (JCP). Interpersonal Scale T-scores increased at rates ranging from 20% (DSF) and, for the IPP and SHY Scales, did not change by ≥ 5 T-score points at 100% random response insertion. PSY-5 Scale T-scores increased at rates ranging from 10% (PSYC-r) and, for the AGGR-r Scale, did not change by ≥ 5 T-score points at 100% random insertion. Therefore, PSYC-r was the scale most susceptible to random response insertion, as evidenced by a mean scale change of ≥ 5 T-score points at 10% simulated response insertion. At 20% response insertion, the GIC, NUC, SUI, AXY, BRF, and DSF Scales evidenced a T-score change of this magnitude. As with the previously presented normative sample results, the VRIN-r mean scores at 10% and 20% random response insertion were less than 80T. Therefore, VRIN-r would not identify a portion of these cases. Specifically, approximately 8.1% of cases would reach a VRIN-r T-Score of ≥ 80T at 20% response insertion (Handel et al., 2010).

In addition to examining increases in SP and PSY-5 Scale mean T-scores, 95% confidence intervals and SEM values were calculated. Confidence intervals for mean T-

scores between levels of response insertion described above (e.g., 10% random response insertion for PSYC-r) and baseline (i.e., 0% response insertion) were non-overlapping. Further, the confidence intervals at 100% random response insertion had limited overlap as compared to baseline conditions. However, as discussed previously, mean T-scores can still be statistically significantly different from one another even in the presence of overlapping confidence intervals (Wolfe & Hanley, 2002).

Regarding the SEM analyses, SEM values spanned from 4 T-score points (SAV and SHY) to 10 T-score points (SUI). The magnitude of deviation between mean scale T-scores at 100% response insertion and baseline ranged from 0 to 6. A change by an SEM value of at least one was observed for scales in each SP and PSY-5 Scale cluster at the following levels of random response insertion: 30% (GIC and NUC) to 90% (MLS) for the Somatic/Cognitive Scales; 40% (SUI) and the SFD Scale did not deviate by one SEM for the RCd-Associated Internalizing SP Scales; 30% (BRF) and the STW and MSF Scales did not deviate by one SEM for the RC7-Associated Internalizing Scales; 50% (AGG) and the JCP and ACT Scales did not deviate by one SEM for the Externalizing Scales; 30% (DSF) and the IPP and SHY Scales did not deviate by one SEM for the Interpersonal Scales; and 20% (PSYC-r) and the AGGR-r and DISC-r Scales did not deviate by one SEM for the PSY-5 Scales. Therefore, the PSYC-r Scale was the most susceptible to random response insertion, as evidenced by a change of one SEM at 20% insertion. At 30%, the GIC, NUC, BRF, and DSF Scales evidenced a change of this magnitude. As noted above, because the mean VRIN-r T-Scores at these levels of response insertion were less than 80T, VRIN-r would not identify a portion of these cases.

Fixed response insertion. The results of increasing degrees of simulated fixed acquiescent and counter-acquiescent responding on SP and PSY-5 Scale scores are presented in Tables 7 through 12 (previously displayed). Under conditions of simulated degrees of fixed acquiescent (i.e., "true") responding, 23 of the 28 SP and PSY-5 Scales increased monotonically. Four scales, which included MLS, IPP, SAV, and INTR-r, decreased monotonically. MSF, the remaining scale, increased from baseline to 100% insertion, but did not increase consistently across each level of insertion. Under conditions of simulated increasing degrees of fixed counter-acquiescent (i.e., "false") responding, 11 of 28 SP and PSY-5 Scales increased monotonically, while 14 Scales decreased monotonically. These scales included SUI, SFD, NFC, STW, AXY, JCP, SUB, AGG, ACT, SHY, AGGR-r, PSYC-r, DISC-r, and NEGE-r. One scale, HLP, increased from baseline to 100% insertion, but did not increase consistently across each level of insertion; the scales ANP and FML decreased in a similar pattern. As with the VRIN-r inpatient analyses, SP and PSY-5 Scale T-scores differed in their susceptibility to fixed acquiescent and counter-acquiescent responding. These differences are discussed in the context of the same three indicators discussed above.

Acquiescent response insertion. For the fixed acquiescent response insertion analyses, a change of ≥ 5 T-score points was observed for the following SP and PSY-5 Scales. Somatic/Cognitive Scale T-scores increased by ≥ 5 T-score points at insertion rates ranging from 10% (GIC) to 50% (NUC); a decrease of ≥ 5 T-score points was observed at 60% (MLS). Internalizing Scale T-scores increased at rates ranging from 10% (SUI, AXY, and BRF) and the MSF Scale did not change by ≥ 5 T-score points. Each of the four Externalizing Scale T-scores (JCP, SUB, AGG, and ACT) increased by

≥ 5 T-score points at 20% response insertion. Interpersonal Scale T-scores increased at rates ranging from 10% (DSF) to 40% (SHY); a decrease of ≥ 5 T-score points was observed for 50% and 70% response insertion for IPP and SAV, respectively. PSY-5 Scale T-scores increased at response rates ranging from 10% (PSYC-r) to 30% (AGGR-r and NEGE-r); a decrease of ≥ 5 T-score points was observed at 30% (INTR-r). Therefore, the GIC, SUI, AXY, BRF, DSF, and PSYC-r were the scales most susceptible to fixed acquiescent response insertion, as evidenced by mean scale T-score changes of ≥ 5 T-score points at 10% simulated response insertion. At 20% response insertion, the COG, HLP, NFC, JCP, SUB, AGG, and ACT evidenced a T-score change of this magnitude. The TRIN-r mean scores at 10% and 20% fixed acquiescent response insertion were less than 80T. Therefore, TRIN-r would not identify a portion of these cases. Specifically, approximately 29.4% of cases would reach a TRIN-r T-Score of ≥ 80T at 20% acquiescent response insertion (Handel et al., 2010).

In addition to examining SP and PSY-5 Scale mean T-scores, 95% confidence intervals and SEM values were calculated. Confidence intervals for mean T-scores between levels of response insertion described above (e.g., 10% fixed acquiescent response insertion for GIC) and baseline (i.e., 0% response insertion) were non-overlapping. Further, the confidence intervals for mean T-scores at 90% fixed acquiescent response insertion had limited overlap as compared to baseline conditions. However, as noted above, mean T-scores can still be statistically significantly different from one another even in the presence of overlapping confidence intervals (Wolfe & Hanley, 2002). As with the normative sample analyses, confidence intervals at 100% response insertion could not be calculated for this condition because the standard

deviation becomes zero when all responses become constant (i.e., all items were changed to *true* responses).

Regarding the SEM analyses, SEM values spanned from 4 T-score points (SAV and SHY) to 10 T-score points (SUI). The magnitude of deviation between mean scale T-scores at 100% response insertion and baseline ranged from 1 to 13. An SEM value of one was observed for scales in each SP and PSY-5 Scale cluster at the following levels of fixed acquiescent response insertion: 20% (GIC and COG) to 70% (MLS and NUC) for the Somatic/Cognitive Scales; 20% (SUI) to 40% (SFD) for the RCd-Associated Internalizing Scales; 20% (AXY and BRF) and the MSF Scale did not deviate by one SEM for the RC7- Associated Internalizing Scales; 20% (AGG) to 30% (JCP, SUB, and ACT) for the Externalizing Scales; 20% (DSF) to 70% (SAV) for the Interpersonal Scales; and 10% (PSYC-r) to 30% (AGGR-r, DISC-r, NEGE-r, and INTR-r) for the PSY-5 Scales. Therefore, the PSYC-r Scale was the most susceptible to fixed acquiescent response insertion, as evidenced by a change of one SEM at 10% insertion. At 20%, the GIC, COG, SUI, AXY, BRF, AGG, and DSF Scales evidenced a change of this magnitude. Mean TRIN-r T-Scores at these levels of response insertion were less than 80T. Therefore, TRIN-r would not identify a portion of these cases.

Counter-acquiescent response insertion. For the fixed counter-acquiescent response insertion analyses, a change of ≥ 5 T-score points was observed for the following SP and PSY-5 Scales. Somatic/Cognitive Scale T-scores increased by ≥ 5 T-score points at insertion rates ranging from 20% (MLS and NUC) and did not change by this amount for the COG Scale. None of the Internalizing Scale T-Scores increased by ≥ 5 T-score points; a decrease of this amount was observed at 40% (SUI, SFD, and NFC) and did not

decrease by ≥ 5 T-score points for ANP. None of the Externalizing Scale T-Scores increased by ≥ 5 T-score points; a decrease of this amount was observed at rates ranging from 30% (JCP) to 80% (SUB). Interpersonal Scale T-scores increased by ≥ 5 T-score points at rates ranging from 30% (SAV and IPP) to 60% (DSF); a decrease by ≥ 5 T-score points was observed at rates ranging from 60% (SHY) and, for the FML scale, did not decrease this amount. PSY-5 Scale T-scores increased by ≥ 5 T-score points at 20% (INTR-r); a decrease by ≥ 5 T-score points was observed at rates ranging from 30% (AGGR-r) and, for the PSYC-r Scale, did not decrease by this amount. Therefore, the MLS, NUC, and INTR-r scales were the most susceptible to fixed counter-acquiescent response insertion, as evidenced by mean scale T-score changes of ≥ 5 T-score points at 20% simulated response insertion. At 30% response insertion, the JCP, IPP, SAV, and AGGR-r Scales evidenced a T-score change of this magnitude. TRIN-r mean scores at 20% and 30% fixed acquiescent response insertion were less than 80T. Therefore, TRIN-r would not identify a portion of these cases. Specifically, approximately 36.0% of cases would reach a TRIN-r T-Score of ≥ 80T at 30% response insertion (Handel et al., 2010).

In addition to examining SP and PSY-5 Scale mean T-scores, 95% confidence intervals and SEM values were calculated. Confidence intervals for mean T-scores between levels of response insertion described above (e.g., 20% fixed counter-acquiescent response insertion for MLS) and baseline (i.e., 0% response insertion) were non-overlapping. Further, the confidence intervals for mean T-scores at 90% fixed counter-acquiescent response insertion had limited overlap as compared to baseline conditions. However, as noted above, mean T-scores can still be statistically significantly different from one another even in the presence of overlapping confidence intervals

(Wolfe & Hanley, 2002). As with the normative sample analyses, confidence intervals at 100% response insertion could not be calculated for this condition because the standard deviation becomes zero when all responses become constant (i.e., all items were changed to *false* responses).

Regarding the SEM analyses, SEM values spanned from 4 T-score points (SAV and SHY) to 10 T-score points (SUI). The magnitude of deviation between mean scale T-scores at 100% response insertion and baseline ranged from 0 to 7. A change by an SEM of at least one was observed for scales in each SP and PSY-5 Scale cluster at the following levels of counter-acquiescent response insertion: 20% (NUC) and the COG Scale did not deviate by one SEM for the Somatic/Cognitive Scales; 40% (NFC) and the HLP Scale did not deviate by one SEM for the RCd-Associated Internalizing Scales; 100% (AXY) and the STW, ANP, BRF, and MSF Scales did not deviate by one SEM for the RC7-Associated Scales; 40% (ACT) and the SUB Scale did not deviate by one SEM for the Externalizing Scales; 30% (IPP and SAV) and the FML Scale did not deviate by one SEM for the Interpersonal Scales; and 20% (INTR-r) and the PSYC-r Scale did not deviate by one SEM for the PSY-5 Scales. Therefore, the NUC and INTR-r Scales were the most susceptible to fixed counter-acquiescent response insertion, as evidenced by a change of one SEM at 20% response insertion. At 30%, the IPP and SAV Scales also evidenced a change of this magnitude. TRIN-r T-Scores at these levels of response insertion were less than 80T. Therefore, TRIN-r would not identify a portion of these cases.

Validity coefficient analyses. Results from the validity analyses conducted on the inpatient sample are presented in this section. In examination of Hypothesis 2a, the

effects of increasing degrees of simulated random, fixed acquiescent, and fixed counter-acquiescent responding on convergent validity coefficients are reviewed. In examination of Hypothesis 2b, the results from MMR analyses are presented.

Prior to running the validity coefficient and MMR analyses, eight proposed pairings of SP/PSY-5 Scales and BPRS items were rationally selected based on their perceived relationship with each other. These pairings are presented with the SP or PSY-5 Scale listed first and the BPRS item name and number presented second. The proposed pairings were as follows: MLS – Somatic Concern (BPRS 1), AXY – Anxiety (BPRS 2), PSYC-r – Conceptual Disorganization (BPRS 4), ACT – Grandiosity (BPRS 8), HLP – Depressive Mood (BPRS 9), AGG – Hostility (BPRS 10), PSYC-r – Hallucinatory Behavior (BPRS 12), and PSYC-r – Unusual Thought Content (BPRS 15). After calculating correlations for these and other pairs of SP/PSY-5 Scales and BPRS items, the proposed list of eight pairings was changed to include a total of 12. Pairs from the proposed list were removed if they did not evidence relatively strong initial correlations (Pearson $r \geq .20$); pairs were added to the list if they (a) evidenced relatively strong initial correlations and (b) were rationally related. This process resulted in the following pairings identified for analysis: MLS – Somatic Concern (BPRS 1), GIC – Somatic Concern (BRPS 1), HPC – Somatic Concern (BPRS 1), HLP – Depressive Mood (BPRS 9), STW – Anxiety (BPRS 2), AXY – Anxiety (BPRS 2), IPP – Emotional Withdrawal (BPRS 3), SAV – Emotional Withdrawal (BPRS 3), PSYC-r – Conceptual Disorganization (BPRS 4), PSYC-r – Unusual Thought Content (BPRS 15), INTR-r – Emotional Withdrawal (BPRS 3), and INTR-r - Depressive Mood (BPRS 9).

Validity coefficient degradation. To quantify the effects of increasing degrees of random and fixed responding on the association between inpatient SP/PSY-5 Scales and BPRS items, convergent validity coefficients were calculated. These results are presented in Table 13. The numbers in each cell are Pearson *r*-values, which were calculated by correlating the 12 pairings of raw SP/PSY-5 Scale scores with BPRS item raw scores. SP and PSY-5 Scale abbreviated names are paired with alphabetical coefficients, which represent relevant BPRS constructs. These constructs are described at the bottom of the table. Initial correlations ranged from $r = .19$ (IPP – Emotional Withdrawal and PSYC-r – Conceptual Disorganization) to $r = .42$ (MLS – Somatic Concern).

Table 18

The Effects of Increasing Degrees of Variable, Fixed Acquiescent, and Fixed Counter-Acquiescent Response Insertion on Correlations Between SP/PSY-5 Scales and Brief Psychiatric Rating Scale Items.

Response insertion percentage	MLS[a]	GIC[a]	HPC[a]	HLP[b]	STW[c]	AXY[c]	IPP[d]	SAV[d]	PSYC-r[e/f]	INTR-r[d/b]
0% (R)	.42	.22	.29	.28	.23	.24	.19	.25	.19/.31	.30/.26
0% (T or F)	.43	.24	.29	.25	.26	.23	.20	.23	.21/.32	.28/.24
10% (R)	.42	.16	.30	.26	.20	.23	.17	.21	.19/.31	.25/.22
10% (T)	.40	.22	.29	.22	.24	.23	.20	.23	.20/.33	.30/.25
10% (F)	.38	.23	.30	.26	.23	.21	.21	.23	.21/.30	.29/.24
20% (R)	.35	.16	.23	.23	.15	.16	.18	.24	.14/.25	.26/.26
20% (T)	.41	.13	.27	.19	.22	.11	.17	.23	.14/.23	.24/.23
20% (F)	.41	.17	.21	.21	.25	.26	.20	.26	.23/.32	.33/.22
30% (R)	.26	.10	.19	.24	.19	.20	.11	.19	.17/.24	.28/.24
30% (T)	.40	.13	.26	.19	.14	.13	.17	.21	.17/.25	-.01/.24

Table 18 Continued

Response insertion percentage	MLS[a]	GIC[a]	HPC[a]	HLP[b]	STW[c]	AXY[c]	IPP[d]	SAV[d]	PSYC-r[e/f]	INTR-r[d/b]
30% (F)	.37	.26	.21	.30	.20	.17	.15	.22	.25/.35	.25/.23
40% (R)	.30	.12	.20	.17	.23	.04	.13	.20	.15/.21	.19/.16
40% (T)	.34	.06	.30	.17	.13	.20	.15	.16	.01/.11	.27/.23
40% (F)	.28	.16	.20	.18	.25	.20	.12	.24	.18/.28	.25/.19
50% (R)	.28	.12	.09	.22	.14	.07	.13	.19	.11/.16	.24/.14
50% (T)	.36	.04	.13	.08	.31	.00	.16	.29	.14/.21	.30/.30
50% (F)	.34	.17	.21	.24	.25	.23	.20	.18	.18/.26	.18/.19
60% (R)	.21	.05	.21	.19	.14	.10	.10	.11	.13/.24	.16/.19
60% (T)	.31	.17	.02	.08	.12	.08	.14	.16	.02/.06	.23/.23
60% (F)	.15	.07	.17	.11	.10	.10	.15	.15	.17/.26	.19/.18
70% (R)	.11	.03	.20	.03	.20	.14	.05	.13	.12/.09	.11/.18
70% (T)	.23	-.14	.22	.07	.12	.09	.14	.19	.15/.06	.22/.17
70% (F)	.25	.11	.17	.14	.05	.09	.05	.21	.14/.17	.19/.25

Table 18 Continued

Response insertion percentage	MLS[a]	GIC[a]	HPC[a]	HLP[b]	STW[c]	AXY[c]	IPP[d]	SAV[d]	PSYC-r[e/f]	INTR-r[d/b]
80% (R)	.10	.01	.01	.07	.11	-.04	-.03	.11	.04/.02	.13/.10
80% (T)	.08	.11	.06	.14	.07	.09	.08	.18	.03/.01	.23/.21
80% (F)	.15	.14	.20	.12	.09	.12	.08	.14	.10/.22	.17/.20
90% (R)	.01	.03	.16	.05	.06	.05	.03	.16	-.02/.00	.11/-.05
90% (T)	.09	.06	-.02	.05	.13	.12	.13	.10	.02/.13	.18/.08
90% (F)	.16	.10	.05	.08	.08	.09	.07	.10	.10/.12	.04/.03
100% (R)	-.04	.07	-.03	-.11	-.12	.01	-.07	-.21	.03/.01	.04/-.05
100% (T)	--	--	--	--	--	--	--	--	--/--	--/--
100% (F)	--	--	--	--	--	--	--	--	--/--	--/--

Note. Variable response insertion n's range from 232 to 235. Fixed acquiescent and fixed counter-acquiescent insertion n's range from 230 to 232. MLS = Malaise; GIC = Gastrointestinal Complaints; HPC = Head Pain

Table 18 Continued

Complaints; HLP = Helplessness/Hopelessness; STW = Stress/Worry; AXY = Anxiety; IPP = Interpersonal Passivity; SAV = Social Avoidance; PSYC-r = Psychoticism-Revised; INTR-r = Introversion/Low Positive Emotions. R = random response insertion; T = fixed acquiescent response insertion; F = counter-acquiescent response insertion.

[a] Somatic Concern, [b] Depressive Mood, [c] Anxiety, [d] Emotional Withdrawal, [e] Conceptual Disorganization, and [f] Unusual Thought Content.

The PSYC-r and INTR-r PSY-5 Scales have two BPRS criterion variables with validity coefficients before and after the /. -- = Correlation could not be calculated because T-scores are constant at 100% fixed acquiescent and fixed counter-acquiescent response insertion.

The inpatient sample (Archer, Griffin, & Aiduk, 1995) is the basis for all confidence interval calculations.

Overall, increasing degrees of simulated random, fixed acquiescent, and fixed counter-acquiescent responding resulted in the degradation of convergent validity coefficients. To determine the magnitude of this degradation across levels of simulated insertion, Pearson *r*-values were squared and then compared with their baseline values (Handel et al., 2010). For simulated random responding, validity coefficients were largely unchanged at 20% response insertion for 11 of the 12 pairings. The exception was the MLS – Somatic Concern pairing, which evidenced a 6% loss in variance accounted for as compared to baseline. At 100% random response insertion, all variable pairings with the exception of IPP – Emotional Withdrawal (3% loss in variance accounted for) and SAV – Emotional Withdrawal (2% loss in variance accounted for) experienced variance losses ranging from 4-18%.

For simulated fixed acquiescent responding, validity coefficients were largely unchanged at 20% for 9 of the 12 pairings. The exceptions were the following pairings: GIC – Somatic Concern (4% loss in variance accounted for), AXY – Anxiety (4% loss in variance accounted for), and PSYC-r – Unusual Thought Content (5% loss in variance accounted for). At 90% fixed acquiescent response insertion, all variable pairings with the exception of IPP – Emotional Withdrawal (2% loss in variance accounted for) experienced variance accounted for losses ranging from 4-17%.

Finally, and for simulated fixed counter-acquiescent responding, validity coefficients were relatively unchanged at 40% for 10 of the 12 pairings. The exceptions were the following pairings: MLS – Somatic Concern (4% loss in variance accounted for) and HPC – Somatic Concern (4% loss in variance accounted for). At 90% fixed counter-acquiescent response insertion, all variable pairings with the exception of PSYC-r –

Conceptual Disorganization (3% loss in variance accounted for) experienced variance losses ranging from 4-15%.

MMR analyses. To examine how random and fixed responding moderate the relationship between SP/PSY-5 Scales and BPRS variables, a total of 12 MMR analyses, each containing three separate regression equations, were conducted. However, prior to conducting the MMR analyses, regression assumptions were checked. It should be noted that some minor violations of these assumptions were expected given that the dependent variables (i.e., BPRS items) are based on a 7-point Likert scale and one of the assumptions of regression is that the dependent variable is unbounded and continuous (Field, 2009; Laerd Statistics, 2015). Violations of these assumptions in the context of interpreting results will be addressed in the Discussion section.

For the random response insertion condition, Durbin-Watson statistic values for each SP/PSY-5 Scale and BPRS pairing except HPC-BPRS1 were close to 2, the value recommended to indicate independence of observations (Laerd Statistics, 2015). Eight of the 12 pairings had VIF values < 10, indicating no concerns about multicollinearity; three of the pairings (PSYC-r – BPRS4, PSYC-r – BPRS15, and INTR-r – BPRS3) had values > 10. This was likely the result of the interaction term that was derived from uncentered raw scores. However, as discussed previously, regression results using centered versus uncentered interaction terms are equivalent (Kromrey & Foster-Johnson, 1998). There were no outliers identified among these pairings, as evidenced by leverage values < 0.2 and Cook's Distance values < 1 (Laerd Statistics, 2015). The assumption of homoscedasticity was likely violated for each of the pairings, as evidenced by visual inspection of plots graphing studentized residuals versus unstandardized predicted values.

Finally, half of the pairings evidenced normal distribution of errors, which was determined by examining Normal Q-Q plots. The remaining half evidenced some distortion; however, regression is relatively robust in this scenario, as the likelihood of finding significant results remains high in the face of this distortion (Minitab, 2015).

For the fixed acquiescent insertion condition, Durbin-Watson statistic values for 10 of the 12 SP/PSY-5 Scale and BPRS pairings were close to 2, the value recommended to indicate independence of observations (Laerd Statistics, 2015). The GIC – BPRS1 and HPC – BPRS1 pairings had values < 1. Each of the 12 pairings had VIF values > 10, indicating the presence of multicollinearity. As noted above, however, this was likely the result of the uncentered interaction term and does not pose a significant threat to regression results (Kromrey & Foster-Johnson, 1998). There were no outliers identified among these pairings, as evidenced by leverage values < 0.2 and Cook's Distance values < 1 (Laerd Statistics, 2015). The assumption of homoscedasticity was likely violated for each of the pairings, as evidenced by visual inspection of plots graphing studentized residuals versus unstandardized predicted values. Finally, 10 of the 12 pairings evidenced normal distribution of errors, which was determined by examining Normal Q-Q plots. The remaining two (GIC – BPRS1 and PSYC-r – BPRS15) evidenced some distortion; however, as discussed above, regression is relatively robust to this violation (Minitab, 2015).

For the fixed counter-acquiescent insertion condition, Durbin-Watson statistic values for 10 of the 12 SP/PSY-5 Scale and BPRS pairings were close to 2, the value recommended to indicate independence of observations (Laerd Statistics, 2015). The GIC – BPRS1 and HPC – BPRS1 pairings had values < 1. All but one (GIC – BPRS1) of the

12 pairings had VIF values > 10, indicating the presence of multicollinearity. As noted for the random and fixed acquiescent conditions, however, this does not represent an area of concern (Kromrey & Foster-Johnson, 1998). There were no outliers identified among these pairings, as evidenced by leverage values < 0.2 and Cook's Distance values < 1 (Laerd Statistics, 2015). The assumption of homoscedasticity was likely violated for each of the pairings, as evidenced by visual inspection of plots graphing studentized residuals versus unstandardized predicted values. Finally, 8 of the 12 pairings evidenced normal distribution of errors, which was determined by examining Normal Q-Q plots. The remaining four (MLS – BPRS1, GIC – BPRS1, PSYC-r – BPRS4, and PSYC-r – BPRS15) evidenced some distortion; however, as discussed above, regression is relatively robust to this violation (Minitab, 2015).

Results from the MMR analyses for the random, fixed acquiescent, and fixed counter-acquiescent response insertion conditions are presented in Tables 14 through 16, respectively. For the random response insertion condition, predictors of BPRS item scores included raw SP/PSY-5 Scale scores, VRIN-r raw scores, and the uncentered cross-product of a SP/PSY-5 Scale and VRIN-r raw scores. Of the 12 MMR analyses completed, one (PSYC-r and BPRS15) evidenced significant slope differences and two (SAV – BPRS3 and INTR-r – BPRS9) had significant intercept differences; the remaining regressions did not evidence any significant moderation effects. The median unadjusted delta R^2 (ΔR^2), which measures the change of variance explained by the addition of the moderating variable, was small (Median = 3%). Examination of the significant slope difference for the PSYC-r and BPRS15 pairing indicated that at higher levels of VRIN-r, increases in PSYC-r scores were more strongly related to increases in

BPRS15 scores as compared to lower levels of VRIN-r. Examination of the intercept differences for the two pairings described above indicated that while VRIN-r scores did not influence the relationship between, for example, SAV and BPRS3 scores, protocols high on VRIN-r scores had higher levels of BPRS3 scores at every level of VRIN-r than their counterparts with lower levels of VRIN-r.

Table 19

The Results of MMR Analyses on SP/PSY-5 and BPRS Variable Pairings Under Conditions of Variable Response Insertion.

	Regression 1 (Overall Moderation)	Regression 2 (Slope Differences)		Regression 3 (Intercept Differences)		ΔR^2
Variable Pairing	(*p*)	(*p*)	B	(*p*)	B	
MLS – BPRS1	0.100	--	--	--	--	0.02
GIC – BPRS1	0.136	--	--	--	--	0.02
HPC – BPRS1	0.630	--	--	--	--	0.00
HLP – BPRS9	0.419	--	--	--	--	0.01
STW – BPRS2	0.195	--	--	--	--	0.01
AXY – BRPS2	0.197	--	--	--	--	0.01
IPP – BPRS3	0.118	--	--	--	--	0.02
SAV – BPRS3	0.024*	--	--	0.012*	-0.06	0.03
PSYC-r – BPRS4	0.198	--	--	--	--	0.01

Table 19 Continued

	Regression 1 (Overall Moderation)	Regression 2 (Slope Differences)		Regression 3 (Intercept Differences)		ΔR^2
Variable Pairing	(*p*)	(*p*)	B	(*p*)	B	
PSYC-r – BPRS15	0.006**	0.002**	-.015	--	--	0.04
INTR-r – BPRS3	0.036*	--	--	0.012**	-0.06	0.03
INTR-r – BPRS9	0.477	--	--	--	--	0.01

Note. n = 277 for the variable response insertion condition. MLS = Malaise; GIC = Gastrointestinal Complaints; HPC = Head Pain Complaints; HLP = Helplessness/Hopelessness; STW = Stress/Worry; AXY = Anxiety; IPP = Interpersonal Passivity; SAV = Social Avoidance; PSYC-r = Psychoticism-Revised; INTR-r = Introversion/Low Positive Emotions-Revised; BPRS1 = Somatic Concern; BPRS2 = Anxiety; BPRS3 = Emotional Withdrawal; BPRS4 = Conceptual Disorganization; BPRS9 = Depressive Mood; BPRS15 = Unusual Thought Content; (*p*) = *p*-value;

Table 19 Continued

B = Unstandardized Regression Coefficient; and ΔR^2 is the change in R^2 from the regression model with only the SP/PSY-5 Scale entered versus the model with the SP/PSY-5 Scale, moderator (VRIN-r), and interaction term added. -- = non-significant values.

* $p < .05$, ** $p < .01$

Table 20

The Results of MMR Analyses on SP/PSY-5 and BPRS Variable Pairings Under Conditions of Fixed Acquiescent Response Insertion.

	Regression 1 (Overall Moderation)	Regression 2 (Slope Differences)		Regression 3 (Intercept Differences)		ΔR^2
Variable Pairing	(*p*)	(*p*)	B	(*p*)	B	
MLS – BPRS1	0.632	--	--	--	--	0.00
GIC – BPRS1	0.012*	--	--	0.011	-0.07	0.04
HPC – BPRS1	0.001**	--	--	0.000***	-0.07	0.06
HLP – BPRS9	0.057	--	--	--	--	0.02
STW – BPRS2	0.085	--	--	--	--	0.02
AXY – BRPS2	0.114	--	--	--	--	0.02
IPP – BPRS3	0.040	0.012*	0.03	--	--	0.03
SAV – BPRS3	0.004**	0.001*	0.05	--	--	0.05
PSYC-r – BPRS4	0.518	--	--	--	--	0.01

Table 20 Continued

	Regression 1 (Overall Moderation)	Regression 2 (Slope Differences)		Regression 3 (Intercept Differences)		ΔR^2
Variable Pairing	(*p*)	(*p*)	B	(*p*)	B	
PSYC-r – BPRS15	0.003**	0.002**	-0.01	--	--	0.05
INTR-r – BPRS3	0.000***	0.000***	0.03	--	--	0.07
INTR-r – BPRS9	0.015*	--	--	0.025*	0.05	0.04

Note. n = 275 for the fixed acquiescent response insertion condition. MLS = Malaise; GIC = Gastrointestinal Complaints; HPC = Head Pain Complaints; HLP = Helplessness/Hopelessness; STW = Stress/Worry; AXY = Anxiety; IPP = Interpersonal Passivity; SAV = Social Avoidance; PSYC-r = Psychoticism-Revised; INTR-r = Introversion/Low Positive Emotions-Revised; BPRS1 = Somatic Concern; BPRS2 = Anxiety; BPRS3 = Emotional Withdrawal; BPRS4 = Conceptual Disorganization; BPRS9 = Depressive Mood; BPRS15 = Unusual Thought

Table 20 Continued

Content; (*p*) = *p*-value; B = Unstandardized Regression Coefficient; and Adjusted ΔR^2 is the change in R^2 from the regression model with only the SP/PSY-5 Scale entered versus the model with the SP/PSY-5 Scale, moderator (TRIN-r), and interaction term added.

-- = non-significant values.

* $p < .05$

** $p < .01$

*** $p < .001$

Table 21

The Results of MMR Analyses on SP/PSY-5 and BPRS Variable Pairings Under Conditions of Fixed Counter-Acquiescent Response Insertion.

	Regression 1 (Overall Moderation)	Regression 2 (Slope Differences)		Regression 3 (Intercept Differences)		ΔR^2
Variable Pairing	(*p*)	(*p*)	B	(*p*)	B	
MLS – BPRS1	0.003*	--	--	0.001**	0.10	0.05
GIC – BPRS1	0.096	--	--	--	--	0.02
HPC – BPRS1	0.101	--	--	--	--	0.01
HLP – BPRS9	0.855	--	--	--	--	0.00
STW – BPRS2	0.617	--	--	--	--	0.00
AXY – BRPS2	0.229	--	--	--	--	0.01
IPP – BPRS3	0.016*	--	--	0.007**	0.09	0.04
SAV – BPRS3	0.008**	--	--	0.002**	0.10	0.04
PSYC-r – BPRS4	0.726	--	--	--	--	0.00

Table 21 Continued

	Regression 1 (Overall Moderation)	Regression 2 (Slope Differences)		Regression 3 (Intercept Differences)		ΔR^2
Variable Pairing	(*p*)	(*p*)	B	(*p*)	B	
PSYC-r – BPRS15	0.438	--	--	--	--	0.01
INTR-r – BPRS3	0.000***	--	--	0.000***	0.20	0.09
INTR-r – BPRS9	0.018*	--	--	0.004**	0.10	0.04

Note. n = 275 for the fixed counter-acquiescent response insertion condition. MLS = Malaise; GIC = Gastrointestinal Complaints; HPC = Head Pain Complaints; HLP = Helplessness/Hopelessness; STW = Stress/Worry; AXY = Anxiety; IPP = Interpersonal Passivity; SAV = Social Avoidance; PSYC-r = Psychoticism-Revised; INTR-r = Introversion/Low Positive Emotions-Revised; BPRS1 = Somatic Concern; BPRS2 = Anxiety; BPRS3 = Emotional Withdrawal; BPRS4 = Conceptual Disorganization; BPRS9 = Depressive Mood; BPRS15 = Unusual

Table 21 Continued

Thought Content; (*p*) = *p*-value; B = Unstandardized Regression Coefficient; and Adjusted ΔR^2 is the change in R^2 from the regression model with only the SP/PSY-5 Scale entered versus the model with the SP/PSY-5 Scale, moderator (TRIN-r), and interaction term added.

-- = non-significant values.

* $p < .05$

** $p < .01$

*** $p < .001$

For the fixed acquiescent response insertion condition, predictors of BPRS variables included raw SP/PSY-5 Scale scores, TRIN-r raw scores, and their uncentered cross-product. Of the 12 MMR analyses completed, four (IPP – BPRS3, SAV – BPRS3, PSYC-r – BPRS15, and INTR-r – BPRS3) evidenced significant slope differences and three (GIC – BPRS1, HPC – BPRS1, and INTR-r – BPRS9) had significant intercept differences; the remaining regressions did not evidence any significant moderation effects. The median ΔR^2 was small (Median = 5%). Examination of the slope differences for the IPP-BPRS3, SAV-BPRS3, and INTR-r-BPRS3 pairings indicated that at higher levels of TRIN-r, increases in the respective SP/PSY-5 Scale were more strongly related to decreases in BPRS3 scores as compared to lower TRIN-r scores. For the remaining pairing (PSYC-r – BPRS15), examination of the slope differences indicated the opposite effect (i.e., at higher levels of TRIN-r, increases in PSYC-r scores were more strongly related to increases in BPRS15 scores as compared to lower levels of TRIN-r). Similarly, examination of the intercept differences for the GIC – BPRS1 and HPC – BPRS1 pairings indicated that while TRIN-r scores did not influence the relationship between, for example, GIC and BPRS1 scores, protocols high on TRIN-r scores had higher levels of BPRS1 scores at every level of TRIN-r than their counterparts with lower levels of TRIN-r. The opposite effect was observed when examining the intercept differences for the INTR-r – BPRS3 pairing.

Finally, for the fixed counter-acquiescent response insertion condition, predictors of BPRS variables also included raw SP/PSY-5 Scale scores, TRIN-r raw scores, and their uncentered cross-product. Of the 12 MMR analyses completed, five (MLS – BPRS1, IPP – BPRS3, SAV – BPRS3, INTR-r – BPRS3, and INTR-r – BPRS9) had significant

intercept differences; the remaining regressions did not evidence any significant moderation effects. The median adjusted ΔR^2 was small (Median = 4%). Examination of these five intercept differences indicated that while TRIN-r scores did not influence the relationship between, for example, MLS and BPRS1 scores, protocols high on TRIN-r scores had higher levels of BPRS1 scores at every level of TRIN-r than their counterparts with lower levels of TRIN-r.

CHAPTER VII

DISCUSSION

The MMPI-2-RF is a widely used and extensively researched instrument that provides researchers and clinicians with a broadband assessment of psychopathology and personality (Ben-Porath, 2012). Of its 51 total Validity and Substantive Scales, 23 SP Scales were designed to augment the RC Scales in the assessment of psychopathology, while five PSY-5 Scales were included to provide a dimensional model of personality pathology. However, the self-report format of the MMPI-2-RF suggests that interpretation of these scales and the clinical recommendations that follow are vulnerable to invalid response styles. While the bulk of existing research has examined the deleterious effects of content-based invalid responding (e.g., overreporting) on information provided by the MMPI-2-RF, less focus has been devoted to the effects of non-content-based invalid responding (e.g., random responding). Therefore, the overall purpose of this dissertation was to examine how simulated non-content-based invalid responding, specifically random (as represented by VRIN-r) and fixed acquiescent and counter-acquiescent (as represented by TRIN-r), affects score interpretation and criterion validity for the 28 SP and PSY-5 Scales.

Inconsistent Responding and SP and PSY-5 Scale Means

The primary focus of this dissertation was based on a design from a previous study by Handel et al. (2010). These researchers analyzed the negative impact of increasing degrees of simulated random and fixed responding on mean T-scores of the RC Scales. Similarly, research by Dragon (2012) examined the effects of varying degrees of random responding on the interpretation of H-O, RC, and SP Scales. These authors and others

(e.g., Burchett & Ben-Porath, 2010) called for an extension of these analyses into the SP and PSY-5 Scales. Therefore, my dissertation primarily sought to examine how increasing degrees of random and fixed response insertion would impact mean T-scores, and therefore scale interpretation, for the 28 SP and PSY-5 Scales. In accordance with the results from Handel and colleagues and Dragon, it was hypothesized that increasing degrees of random and fixed response insertion would increase mean T-scores for a majority of these scales. Further, it was proposed that differences in the effects of non-content-based invalid responding on mean scale T-scores would vary based on item keying for each scale (e.g., scales with most or all items keyed *true* would increase more rapidly with simulated acquiescent responding as compared to scales with most or all items keyed *false*). To examine this primary aim, a computer simulation procedure used by Handel and colleagues was used to insert increasing degrees of random and fixed responding into protocols from the nongendered MMPI-2-RF normative sample (N = 2,276; Ben-Porath & Forbey, 2003) and a sample of psychiatric inpatients (N = 704; Archer et al., 1995; Handel & Archer, 2008). Three measures were used to examine the magnitude of mean T-score distortion caused by this experimental manipulation. These included noting mean T-scores that changed by ≥ 5 T-score points (Ben-Porath, 2012), calculating 95% confidence intervals for each mean T-score, and determining the multiples of SEMs an experimental mean T-score (e.g., mean T-score at 40% random response insertion) deviated from the baseline (i.e., 0% response insertion) condition.

Results from this set of analyses supported two primary conclusions, both of which will be elaborated upon below. First, increasing degrees of random and fixed responding resulted in significant mean T-score distortions for most of the 28 SP and PSY-5 Scales.

Second, SP and PSY-5 Scales differed in their susceptibility to non-content-based response insertion. In other words, a response insertion percentage increased, the mean T-scores for some scales changed more quickly than others. As discussed above, this effect was likely the result of each scale's item keying, or combination of *true* and *false* responses counted towards a scale's raw score (Handel et al, 2010), and item endorsement frequencies (Dragon, 2012).

For the normative and inpatient samples increasing degrees of simulated random responding resulted in mean T-scores increases for 27 of the 28 SP and PSY-5 Scales. The exception was the AGGR-r scale for inpatients, which evidenced a small decrease in mean T-scores. As predicted, SP and PSY-5 Scale mean T-scores differed in their susceptibility to score distortion. Specifically, the NUC, GIC, SUI, AXY, BRF, and PSYC-r scales evidenced mean T-score increases, as measured by the three indicators described above, at relatively low levels of random response insertion (i.e., 10 and 20% random responses). This is of particular clinical importance, as the mean VRIN-r T-score at these levels of response insertion was less than 80T. Thus, VRIN-r would not identify a portion of these protocols as invalid, which could result in an interpretive error.

The results are in general accord with those from Handel et al. (2010) and Dragon (2012). Specifically, while both studies found that mean T-scores for the Substantive Scales examined generally increased as a result of random response insertion, there were certain scales more susceptible to this experimental manipulation than others. As discussed briefly above, there are several possible explanations for these differences in susceptibility. First, the relative number of items scored in either the *true* or *false* direction for a scale could relate to these differences. For example, the NUC scale is

comprised of 10 items with three keyed *true* and 7 keyed *false* (Tellegen & Ben-Porath, 2008/2011). Therefore, increasing random response insertion (i.e., *true* and *false* responses) could artificially increase mean scale T-scores.

Second, higher item endorsement frequencies in the normative sample for certain scales could also explain this phenomenon. It should be noted that information about item endorsement frequencies, as reported by Butcher et al. (1989), was drawn from the MMPI-2 normative sample; this information was not available for the MMPI-2-RF (Tellegen & Ben-Porath, 2008/2011). These samples differ in two important ways. First, the MMPI-2 normative sample (N = 2,600) is slightly larger than the MMPI-2-RF normative sample (N = 2,276). Second, item endorsement frequencies for the MMPI-2 were calculated for males and females separately, while the MMPI-2-RF uses T-scores calculated from the combined gender sample. Thus, the following information about item endorsement frequencies should be interpreted with these limitations in mind. Further, it is necessary to report values for male and female participants based on how the information is presented by Butcher and colleagues.

To illustrate how differences in item endorsement frequencies could relate to the differential susceptibility of mean scale T-scores to random response insertion, items in the SUI and SFD Scales will be compared. Item endorsement frequencies in the SUI scale ranged from 2-12% (males) and 2-15% (females) for the normative sample and 22-51% (males) and 24-49% (females) for a comparison psychiatric inpatient sample (Butcher at el., 1989). Alternatively, item endorsement frequencies for the SFD scale ranged from 17-34% (males) and 23-38% (females) in the normative sample and 52-60% (males) and 60-70% (females) in the inpatient sample. Thus, if items with lower

endorsement frequencies were changed to the keyed direction by random response insertion, their impact on scale mean T-scores could be greater than items with higher endorsement frequencies.

Increasing degrees of simulated fixed acquiescent responding resulted in most of the SP and PSY-5 Scale mean T-scores increasing for both the normative and inpatient samples. Four scales (MLS, IPP, SAV, and INTR-r) decreased as simulated responding increased. These four scales are largely or entirely comprised of items keyed in the *false* direction (Ben-Porath & Tellegen, 2008); therefore, it follows that as acquiescent responding (i.e., increasing degrees of *true* responses) increases, mean T-scores for these scales would decrease.

As with the random response insertion analyses, SP and PSY-5 Scales differed in their susceptibility to simulated fixed acquiescent responding. Specifically, the GIC, SUI, AXY, BRF, DSF, AGG, and PSYC-r scales evidenced significant mean T-score increases, as measured by the three indicators described above, at relatively low levels of fixed acquiescent response insertion (i.e., 10 and 20% response insertion). This is of particular clinical importance, as the mean TRIN-r T-score at these levels of response insertion was less than 80T. Thus, TRIN-r would not identify a portion of these protocols as invalid, which could result in an interpretive error. As with the random response insertion analyses, the differential susceptibility of SP and PSY-5 mean scale T-scores to increasing degrees of fixed acquiescent response insertion was likely the result of item keying and differential item endorsement frequencies. For example, the items in the SUI scale have low item endorsement frequencies (Butcher et al., 1989) and are all keyed in the *true* direction (Ben-Porath & Tellegen, 2008).

Finally, increasing degrees of fixed counter-acquiescent responding resulted in SP and PSY-5 mean T-score increases for 18 of 28 Scales in the normative sample and 12 of 28 in the inpatient sample; the remaining scale mean T-scores for each sample decreased. A higher proportion of SP and PSY-5 mean T-score reductions across increasing degrees of fixed counter-acquiescent responding were observed. This was likely the result of a greater number of SP and PSY-5 Scales being comprised of items keyed mostly in the *true* direction (Ben-Porath & Tellegen, 2008).

SP and PSY-5 Scales also differed in their susceptibility to fixed counter-acquiescent response insertion. Specifically, the NUC, MLS, IPP, SAV, and INTR-r scales evidenced significant mean T-score increases, as measured by the three indicators described above, at relatively low levels of fixed counter-acquiescent response insertion (i.e., 10 and 20% response insertion). This is of particular clinical importance, as the mean TRIN-r T-score at these levels of response insertion was less than 80T. Thus, TRIN-r would not identify a portion of these protocols as invalid, which could result in an interpretive error. As with the previous analyses, the differential susceptibility of SP and PSY-5 mean scale T-scores to increasing degrees of fixed acquiescent response insertion was likely the result of item keying. For example, all of the items in the INTR-r scale are keyed in the false direction (Ben-Porath & Tellegen, 2008).

Inconsistent Responding and External Validity

The secondary focus of this dissertation sought two aims. The first aim was also an extension of research conducted by Handel et al. (2010) and Dragon (2012). Specifically, Handel and colleagues quantified the negative impact of random and fixed responding on validity coefficients calculated between the RC Scales and a relevant external criterion

measure, the BPRS. Dragon's study examined the impact of random responding on validity coefficients calculated between the H-O, RC, and SP Scales and a host of external criterion measures. Therefore, this dissertation sought to examine how increasing degrees of random and fixed responding would impact validity coefficients calculated between 12 pairings of SP/PSY-5 Scales and selected BPRS items. In accordance with the results from Handel and colleagues and Dragon, it was hypothesized that convergent validity coefficients for SP and PSY-5 Scales would degrade under conditions of increasing simulated random and fixed responding. Further, these researchers observed the following trend: coefficients were (a) relatively robust to random insertion rates below 30% and (b) evidenced substantial degradations at rates ≥ 30%. Therefore, it was also hypothesized that the same trend would emerge in this study. Finally, given that BPRS data was available for only the psychiatric inpatient sample, it should be noted that validity coefficient analyses were only performed on protocols from this sample. Readers interested in the steps used to examine this aim are referred to the Procedures section.

Overall, increasing degrees of simulated random responding resulted in a pattern of convergent validity coefficient degradation similar to the results reported by Handel et al. (2010) and Dragon (2012). Specifically, validity coefficients for most of the SP/PSY-5 and BPRS variable pairings were relatively robust to random response insertion at rates below 30%. Further, a majority of these pairings evidenced substantial losses in variance accounted for when random response insertion reached 100%. A highly similar pattern of results was observed for the effects of increasing degrees of fixed acquiescent and fixed counter-acquiescent responding on convergent validity coefficients. These results suggest that the SP and PSY-5 Scales are relatively robust to substantial losses of external

validity at lower levels of inconsistent responding. At higher levels, however, external validity is substantially impacted.

The second aim sought to examine how random and fixed responding (as represented by VRIN-r and TRIN-r scores, respectively) moderated the relationship between the same 12 pairings of SP/PSY-5 Scales and BPRS items described above. Burchett (2012) conducted a similar study, and reported that MMPI-2-RF overreporting Validity Scales (e.g., F-r) moderated the relationship between RC Scales and a host relevant external criterion measures. However, specific hypotheses regarding the moderating effects of random and fixed responding were not offered for this study, as this author was not aware of any existing studies that have examined moderation using inconsistency Validity Scales. Readers interested in the steps used to examine this aim are referred to the Procedures section.

Results of the MMR analyses varied based on the type of inconsistent responding simulated. For increasing degrees of random response insertion, three of the 12 variable pairings evidenced significant moderation effects; one slope difference and two intercept differences were found. Increasing degrees of fixed acquiescent response insertion resulted in seven significant moderation effects, with four variable pairings exhibiting significant slope differences and three demonstrating intercept differences. For the fixed counter-acquiescent condition, five variable pairings evidenced significant intercept differences; no slope differences were observed. It should be noted that the median effect size, as measured by delta R^2, for the regressions in which significant moderation effects were found was small (4%). Thus, it is possible that the total number of participants used in these analyses (n = 277 for random response insertion and n = 275 for the fixed

conditions) was not large enough to detect a greater proportion of significant moderation effects. While a substantial portion of protocols were eliminated from the total number (*N* = 704) of the inpatient sample due to validity criteria violations, previous research (Dragon, 2012) has demonstrated that correlations between MMPI-2-RF Substantive Scales and external criterion measures are improved after application of these criteria. Application of validity criteria also follows the basic instructions for test scoring and interpretation procedures (Ben-Porath, 2012). Finally, while some moderation effects were observed, the relatively small proportion of effects detected by these analyses does not detract from the support for using VRIN-r and TRIN-r. Rather, the primary focus of this dissertation discussed previously strongly supports the opposite conclusion.

Implications

The results of this dissertation add to the existing literature base by demonstrating the deleterious effects of random and fixed responding on Substantive Scales of the MMPI-2-RF. Specifically, at lower levels of inconsistent responding, these results suggest that mean scale T-scores can change significantly. While this effect was not found for each of the scales examined, a cluster of scales particularly vulnerable to lower levels of inconsistent responding was identified. This is of particular importance, as interpretation of these scales could be inaccurate even if an examiner determines that validity criteria for a protocol have been met. It is hoped that the results of this dissertation will be incorporated into future test development and interpretive recommendations in two possible ways. First, different validity criteria cutoffs could exist for certain Substantive Scales. Given the complexity of applying differential cutoffs to different scales, however, it is likely this change would be most effectively implemented by integrating these and

future results into the current computerized scoring procedure. Second, and alternatively, item keying for certain Substantive Scales could be modified. For example, the five items from the SUI Scale, which are currently all keyed in the *True* direction, could be modified such that two of five items remain keyed in the *True* direction and the remaining items are keyed in the *False* direction. This would likely reduce the susceptibility of this and other scales to random and fixed responding by limiting the impact of one or two invalid item responses on the mean scale T-score.

Further, at higher levels of inconsistent responding, including those that would result in a protocol being identified as invalid, the results of this dissertation suggest that scales begin to lose external validity. Given that the MMPI-2-RF was designed to accurately assess constructs of psychopathology and personality, these results support the continued use of VRIN-r and TRIN-r in the detection of high degrees of inconsistent responding.

Limitations and Future Directions

This dissertation had several limitations. First, non-content-based invalid responding was introduced to protocols via a computer simulation procedure and not via actual participants (i.e., instructing a sample of participants to engage in inconsistent responding). While the use of this simulation procedure allowed for a graded, detailed, and controlled increase of inconsistent responding, two areas of limitation were identified. First, simulated response insertion may not be an accurate representation of the response styles of actual test-takers (Ben-Porath, 2012; Handel et al., 2010). For example, some test-takers may respond validly to two-thirds of the MMPI-2-RF (i.e., the first approximately 225 items), but respond randomly to the last third (i.e., the last approximately 113 items). Therefore, the effects on scale mean T-scores of this “actual”

response style could possibly differ from the mean T-score changes resulting from the procedures followed in this study (e.g., randomly inserting random responses into 30% of all MMPI-2-RF items from each protocol). Thus, a potentially useful future study would be to replicate this study using actual participants who would be provided instructions to respond in inconsistently. This has been done successfully with other studies (i.e., Burchett, 2012; Dragon, 2012). Second, it was not possible in this study to quantify the number of item responses from each protocol that were changed as a result of response insertion. For example, if the syntax were set such that 30% of items from each protocol were selected for random response insertion, then a portion of these item responses would have remained the same as before the application of the response insertion procedure. In other words, if an item were keyed *True* prior to the procedure and then selected by the syntax for random insertion, it stood only a 50% change of being changed to *False.* Thus, 30% random response insertion might have only resulted in a response change for 15% of items. The same phenomenon was also possible for the fixed acquiescent and counter-acquiescent response conditions. For example, items keyed *True* prior to the application of the procedure would not be changed under increasing degrees of fixed acquiescent response insertion. Thus, future studies may benefit from a more detailed analysis of item response change as a function of random and fixed response insertion.

Second, protocols from the inpatient sample were collected from participants who took the MMPI-2, not the MMPI-2-RF. I transformed the inpatient MMPI-2 protocols into MMPI-2-RF protocols prior to application of the experimental procedure so both samples used in this study would have the same form of this measure. As noted in previous sections of this dissertation, items from both forms were drawn from the same

pool and normative information was based on the same sample. Further, Tellegen & Ben-Porath (2008/2011) reported the results of a comparability study indicating the near equivalency of results obtained from MMPI-2 protocols, MMPI-2-RF protocols, and MMPI-2-RF protocols that were transformed from MMPI-2 protocols (i.e., the procedure followed in this study). However, participants from the Tellegen and Ben-Porath study were college students. Therefore, the generalizability of their results to the current study is unclear; it is possible that results generated from the transformed MMPI-2-RF protocols might have differed slightly from results collected from the same sample of inpatients who took the MMPI-2-RF itself. Thus, future studies may benefit from administration of the MMPI-2-RF itself to inpatient and other non-college student populations.

Third, this study used individual items from one external criterion measure, the BPRS. Further, BPRS data was only available for inpatient sample protocols. This may limit the generalizability of the external validity analyses, and future studies would likely benefit by increasing the number of criterion variables examined. For example, scores on the Beck Depression Inventory-II (BDI; Beck, Steer, & Brown, 1996) could be paired with the scores from MMPI-2-RF Scales HLP and INTR-r to examine the impacts of random and fixed responding on convergent validity coefficients. Similarly, skin conductance levels, a physiological measure of anxiety (Bond, James, & Lader 1974), could be paired with the STW and AXY Scales.

Fourth, and similarly, the use of individual BPRS items in the MMR analyses may have resulted in the observed regression assumption violations. Data on BPRS inter-rater reliability was also not available. Further, and as discussed above, the number of

participants used for the MMR analyses may not have been enough to detect a small effect size. Thus, future studies examining moderation effects between SP/PSY-5 Scales and external criterion measures should be conducted with a sample size estimated using a small predicted effect size.

Summary

The MMPI-2-RF is a widely used instrument in the assessment of psychopathology and personality. The results of this dissertation add to existing literature by demonstrating the conditions under which the interpretation and validity of information provided by MMPI-2-RF scores is compromised. Should these results be replicated in future studies, it is hoped that these findings will contribute to the continued improvement of this test. It is also hoped that future studies address the limitations identified in this dissertation.

REFERENCES

Aiken, L. S., & West, S. G. (1991). Multiple regression: Testing and Interpreting Interactions. Newbury Park, CA: Sage Publications, Inc.

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

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.

Anderson, J.L., Sellbom, M., Bagby, R.M., Quilty, L.C., Veltri, C.O.C., Markon, K.E., & Krueger, R.F. (2012). On the convergence between PSY-5 Domains and PID-5 Domains and Facets: Implications for assessment of DSM-5 personality traits. *Assessment, 20*(3), 286-294. doi: 10.1177/1073191112471141

Anestis, J.C., Gottfried, E.D., & Joiner, T.E. (2015). The utility of MMPI-2-RF Substantive Scales in prediction of negative treatment outcomes in a community mental health center. *Assessment, 22*(1), 23-35. doi: 10.1177/1073191114536771

Aguinis, H., Culpepper, S. A., & Pierce, C. A. (2010). Revival of test bias research in preemployment testing. *Journal of Applied Psychology, 95*(4), 648-680. doi: 10.1037/a0018714

Archer, R. P., Fontaine, J., & McCrae, R. R. (1998). Effects of two MMPI-2 Validity Scales on basic scale relations to external criteria. *Journal of Personality Assessment, 70,* 87-102.

Archer, R. P., Griffin, R., & Aiduk, R. (1995). MMPI-2 clinical correlates for ten common codes. *Journal of Personality Assessment, 65,* 391-407.

Ayearst, L.E., Sellbom, M., Trobst, K.K., & Bagby, R.M. (2013). Evaluating the interpersonal content of the MMPI-2-RF Interpersonal Scales. *Journal of Personality Assessment, 95*(2), 187-196. doi: 10.1080/00223891.2012.730085

Bagby, R. M., Ryder, A. G., Ben Dat, D., Bacchiochi, J., & Parker, J. D. A. (2002). Validation of the dimensional factor structure of the Personality Psychopathology Five in clinical and nonclinical samples. *Journal of Personality Disorders, 16*(4), 304-316. doi: 10.1521/pedi.16.4.304.24128

Beck, A. T., & Steer, R. A. (1993). *BHS: Beck Hopelessness Scale manual.* San Antonio, TX: The Psychological Corporation.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II.* San Antonio, TX: The Psychological Corporation.

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

Ben-Porath, Y. S., & Forbey, J. D. (2003). *Non-gendered norms for the MMPI-2.* Minneapolis, MN: University of Minnesota Press.

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

Berry, D. T. R., Wetter, M. W., Baer, R. A., Larsen, L., Clark, C., & Monroe, K. (1992). MMPI-2 random responding indices: Validation using a self-report methodology. *Psychological Assessment, 4,* 340-345.

Berry, D. T. R., Wetter, M. W., Baer, R. A., Widiger, T. A., Sumpter, J. C., Reynolds, S. K., Hallam, R. A. (1991). Detection of random responding on the MMPI-2: Utility of

F, Back *F*, and *VRIN* scales. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3*, 418-423.

Bond, A. J., James, D. C., & Lader, M. H. (1974). Physiological and psychological measures in anxious patients. *Psychological Medicine, 4*, 364-373.

Brinker, J.K., Chin, Z.H., & Wilkinson, R. (2014). Ruminative thinking style and the MMPI-2-RF. *Personality and Individual Differences, 66,* 102-105. doi: 10.1016/j.paid.2014.03.001

Buechley, R., & Ball, H. (1952). A new test of "validity" for the group MMPI. *Journal of Clinical Psychology, 16*, 299-301.

Burchett, D. L. (2012). *MMPI-2-RF validity scale scores as moderators of substantive scale criterion validity* (Unpublished doctoral dissertation). Kent State University, Kent, OH.

Burchett, D. L., & Ben-Porath, Y. S. (2010). The impact of overreporting on MMPI-2-RF substantive scale validity. *Assessment, 17*(4), 497-516. doi: 10.1177/1073191110378972

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

Butcher, J. N., Hass, G. A., Greene, R. L., & Nelson, L. D. (2015). *Using the MMPI-2 in forensic assessment* [American Psychological Association]. doi: 10.1037/14571-015

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

Camara, W. J., Nathan, J. S., & Puente, A. E. (2000). Psychological test usage: Implications in professional psychology. *Professional Psychology: Research and Practice, 31,* 141-154.

Charter, R. A., & Lopez, M. N. (2003). MMPI-2: Confidence intervals for random responding to the F, F Back, and VRIN scales. *Journal of Clinical Psychology, 59*(9), 985-990. doi: 10.1002/jclp.10183

Clark, M. E., Gironda, R. J., & Young, R. W. (2003). Detection of Back Random Responding: Effectiveness of MMPI-2 and Personality Assessment Inventory validity indices. *Psychological Assessment, 15*(2), 223-234. doi: 10.1037/1040-3590.15.2.223

Cleckley, H. (1982). *The mask of sanity* (5th ed.). St. Louis, MO: Mosby.

Cramer, K. M. (1995). Comparing three new MMPI-2 randomness indices in a novel procedure for random profile derivation. *Journal of Personality Assessment, 65,* 514-520.

Dragon, W.R. (2012). *Effects of random responding on the interpretability of MMPI-2-RF substantive scale scores* (Unpublished doctoral dissertation). Kent State University, Kent, OH.

Faustman, W. O., & Overall, J. E. (1999). Brief Psychiatric Rating Scale. In M. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 791-830). Mahwah, NJ: Erlbaum.

Field, A. (2009). *Discovering statistics using SPSS: Third edition.* Thousand Oaks, CA: Sage Publications Inc.

Forbey, J. D., Lee, T. T. C., & Handel, R. W. (2010). Correlates of the MMPI-2-RF in a college setting. *Psychological Assessment, 22*(4), 737-744. doi: 10.1037/a0020645

Gallen, R. T., & Berry, D. T. R. (1996). Detection of random responding in MMPI-2 protocols. *Assessment, 3,* 171-178.

Gervais, R. O., Ben-Porath, Y. S., & Wygant, D. B. (2009). Empirical correlates and interpretation of the MMPI-2-RF Cognitive Complaints Scale. *Clinical Neuropsychologist, 23*(6), 996-1015. doi: 10.1080/13854040902748249

Glassmire, D. M., Stolberg, R. A., Greene, R. L., & Bongar, B. (2001). The utility of suicide items for assessing suicide potential: Development of a suicide potential scale. *Assessment, 8,* 281-290.

Gottfried, E., Bodell, L., Carbonell, J., & Joiner, T. The clinical utility of the MMPI-2-RF Suicidal/Death Ideation Scale. *Psychological Assessment, 26*(4), 1205-1211. doi: 10.1037/pas0000017

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

Graham, J. R. (2012). *MMPI-2: Assessing personality and psychopathology* (5th ed.). New York, NY: Oxford University Press.

Greene, R. L. (1978). An empirically derived MMPI Carelessness Scale. *Journal of Clinical Psychology, 34,* 407-410.

Greene, R. L. (1979). Response inconsistency on the MMPI: The TR Index. *Journal of Personality Assessment, 43,* 69-77.

Greene, R. L. (1991). *The MMPI-2/MMPI: An interpretive manual.* Boston, MA: Allyn & Bacon.

Haber, J.C., & Baum L.J. (2014). Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF) Scales as predictors of psychiatric diagnoses. *South African Journal of Psychology, 44*(4), 439-453. doi: 10.1177/0081246314532788

Handel, R. W., & Archer, R. P. (2008). An investigation of the psychometric properties of the MMPI-2 Restructured Clinical (RC) Scales with mental health inpatients. *Journal of Personality Assessement, 90*(3), 239-249. doi: 10.1080/00223890701884954

Handel, R. W., Arnau, R. C., Archer, R. P., & Dandy, K. L. (2006). An evaluation of the MMPI-2 and MMPI-A True Response Inconsistency (TRIN) Scales. *Assessment, 13*(1), 98-106. doi: 10.1177/1073191105284453

Handel, R. W., Ben-Porath, Y. S., Tellegen, A., & Archer, R. P. (2007). *Psychometric functioning of the MMPI-2-RF VRIN-r and TRIN-r scales with varying degrees of randomness, acquiescence, and nonacquiescence.* Paper presented at the 42nd annual symposium on Recent Developments in the Use of the MMPI-2/MMPI-2, Fort Lauderdale, FL.

Handel, R. W., Ben-Porath, Y. S., Tellegen, A., & Archer, R. P. (2010). Psychometric functioning of the MMPI-2-RF VRIN-r and TRIN-r scales with varying degrees of randomness, acquiescence, and counter-acquiescence. *Psychological Assessment, 22*(1), 87-95. doi: 10.1037/a0017061

Harkness, A. R. (1992). Fundamental topics in the personality disorders: Candidate trait dimensions from lower regions of the hierarchy. *Psychological Assessment, 4*, 251-259.

Harkness, A. R., Finn, J. A., McNulty, J. L., & Shields, S. M. (2012). The Personality Psychopathology-Five (PSY-5): Recent constructive replication and assessment literature review. *Psychological Assessment, 24*(2), 432-443. doi: 10.1037/a0025830

Harkness, A. R., & McNulty, J. L. (1994). The Personality Psychopathology Five (PSY-5): Issues from the pages of a diagnostic manual instead of a dictionary. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (pp. 291-315). New York, NY: Springer.

Harkness, A. R., & McNulty, J. L. (2006). An overview of personality: The MMPI-2 Personality Psychopathology Five (PSY-5) Scales. In J. N. Butcher (Ed.), *MMPI-2: A practitioner's guide* (pp. 73-97). Washington, DC: American Psychological Association.

Harkness, A. R., & McNulty, J. L. (2007). *Restructured versions of the MMPI-2 Personality Psychopathology Five (PSY-5) Scales.* Paper presented at the meeting of the American Psychological Association, San Francisco, CA.

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

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

Hedlund, J. L., & Vieweg, B. W. (1980). The Brief Psychiatric Rating Scale: A comprehensive review. *Journal of Operational Psychiatry, 11,* 48-62.

Hoelzle, J. B., Nelson, N. W., & Arbisi, P.A. (2012). MMPI-2 and MMPI-2-Restructured Form Validity Scales: Complementary approaches to evaluate response validity. *Psychological Injury and Law, 2,* 174-191. doi: 10.1007/s12207-012-9139-2

Iverson, G. L., & Barton, E. (1999). Interscorer reliability of the MMPI-2: Should TRIN and VRIN be computer scored? *Journal of Clinical Psychology, 55,* 65-69.

Kromrey, J. D., & Foster-Johnson, L. (1998). Mean centering in moderated multiple regression: Much ado about nothing. *Educational and Psychological Measurement, 58,* 42-67.

Krueger, R.F., Derringer, J., Markon, K.E., Watson, D., & Skodol, A.E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine, 42,* 1879-1890.

Laerd Statistics. (2013). *Multiple regression in SPSS.* Retrieved from http://statistics.laerd.com. Accessed April – May 2015.

Lilienfeld, S.O., & Andrews, B.P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal populations. *Journal of Personality Assessment, 66,* 488-524. doi: 10.1207/s15327752jpa6603_3

Malgady, R. G., Rogler, L. H., & Cortes, D. E. (1996). Cultural expression of psychiatric symptoms: Idioms of anger among Puerto Ricans. *Psychological Assessment, 8,* 265-268.

Marek, R.J., Block, A.R., & Ben-Porath, Y.S. (2014). The Minnesota Multiphasic Personality Inventory-2-RF (MMPI-2-RF): Incremental validity in predicting early postoperative outcomes in spine surgery candidates. *Psychological Assessment,* 1-11. doi: 10.1037/pas0000035

Mattson, C., Powers, B., Halfaker, D., Akeson, S., & Ben-Porath, Y. (2012). Predicting Drug Court treatment completion using the MMPI-2-RF. *Psychological Assessment, 24*(4), 937-943. doi: 10.1037/a0028267

Meehl, P.E., & Hathaway, S.R. (1946). The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *Journal of Applied Psychology*, *30*(5), 525-564.

Minitab. (2015). *Multiple regression in the assistant.* Retrieved from http://support.minitab.com/ en-us/minitab/17/technical-papers/. Accessed June 2015.

Nichols, D. S., Greene, R. L., & Schmolck, P. (1989). Criteria for assessing inconsistent patterns of item endorsement on the MMPI: Rationale, development, and empirical trials. *Journal of Clinical Psychology, 45,* 239-250.

Overall, J. E., & Gorham, D. R. (1988). Introduction – the Brief Psychiatric Rating Scale (BPRS): Recent developments in ascertainment and scaling. *Psychological Bulletin, 24,* 97-98.

Paolo, A. M., & Ryan, J. J. (1992). Detection of random response sets on the MMPI-2. *Psychotherapy in Private Practice, 11,* 1-8.

Phillips, T.R., Sellbom, M., Ben-Porath, Y.S., & Patrick, C.J. (2014). Further development and construct validation of the MMPI-2-RF indices of Global Psychopathy, Fearless-Dominance, and Impulsive-Antisociality in a sample of incarcerated women. *Law and Human Behavior*, *38*(1), 34-46. doi: 10.1037/lhb0000040

Pinsoneault, T. B. (2007). Detecting random, partially random, and nonrandom Minnesota Multiphasic Personality Inventory-2 protocols. *Psychological Assessment, 19*(1), 159-164. doi: 10.1037/1040-3590.19.1.159

Rogers, R., Dolmetsch, R., & Cavanaugh, J. L. (1983). Identification of random responders on MMPI Protocols. *Journal of Personality Assessment, 47,* 364-368.

Rouse, S. V. (2007). Using reliability generalization methods to explore measurement error: An illustration using the MMPI-2 PSY-5 scales. *Journal of Personality Assessment, 88*(3), 264-275. doi: 10.1080/00223890701293908

Rouse, S. V., Finger, M. S., & Butcher, J. N. (1999). Advances in clinical personality measurement: An item response theory analysis of the MMPI-2 PSY-5 scales. *Journal of Personality Assessment, 72,* 282-307.

Sellbom, M., Anderson, J.L., and Bagby, M.R. (2013). Assessing DSM-5 Section III personality traits and disorders with the MMPI-2-RF. *Assessment, 20*(6), 709-722. doi: 10.1177/1073191113508808

Sellbom, M., Bagby, R. M., Kushner, S., Quilty, L. C., & Ayearst, L. E. (2012). Diagnostic construct validity of MMPI-2 Restructured Form (MMPI-2-RF) scale scores. *Assessment, 19*(2), 176-186. doi: 10.1177/1073191111428763

Sewell, K. W., & Rogers, R. (1994). Response consistency and the MMPI-2: Development of a simplified screening scale. *Assessment, 1,* 293-299.

Sprock, J. (2000). Invalid response sets in MMPI and MMPI-2 profiles of college students. *Educational and Psychological Measurement, 60,* 956-964.

Tarescavage, A.M., Corey, D.M., & Ben-Porath, Y.S. (2015). Minnesota Multhiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) predictors of police officer problem behavior. *Assessment, 22*(1), 116-132. doi: 10.1177/1073191114534885

Tarescavage, A.M., Luna-Jones, L., & Ben-Porath, Y.S. (2014). Minnesota Multiphasic Personality Inventory-2-Restructed Form (MMPI-2-RF) predictors of violating probations after felonious crimes. *Psychological Assessment, 26*(4),1375-1380. doi: 10.1037/pas0000022

Tarescavage, A.M., Scheman, J., & Ben-Porath (2014). Reliability and validity of the Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) in evaluations of chronic low back pain patients. *Psychological Assessment,* 1-14. doi: 10.1037/pas0000056

Tarescavage, A.M., Wygant, D.B., Boutacoff, L.I., & Ben-Porath, Y.S. (2013). Reliability, validity, and utility of the Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) in assessments of bariatric surgery candidates. *Psychological Assessment, 25*(4), 1179-1194. doi: 10.1037/a0033694

Tellegen, A. (1988). The analysis of consistency in personality assessment. *Journal of Personality, 56*, 621-663.

Tellegen, A. (1995/2003). *Multidimensional Personality Questionnaire-276 (MPQ-276) test booklet.* Minneapolis, MN: University of Minnesota Press.

Tellegen, A., & Ben-Porath, Y. S. (2008/2011). *MMPI-2-RF (Minnesota Multiphasic Personality Inventory-2 Restructured Form) technical manual.* Minneapolis, MN: University of Minnesota Press.

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

van der Heijden, P. T., Egger, J. I. M., Rossi, G. M. P., Grundel, G., & Derksen, J. J. L. (2012). The MMPI-2 Restructured Form and the standard MMPI-2 Clinical Scales in relation to DSM-IV. *European Journal of Psychological Assessment, 29*(3), 182-188. doi: 10.1027/1015-5759/a000140

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology, 33,* 448-457.

Watson, L. C., Quilty, L. C., & Bagby, R. M. (2010). Differentiating Bipolar Disorder from Major Depressive Disorder using the MMPI-2-RF: A receiver operating characteristics (ROC) analysis. *Journal of Psychopathology and Behavioral Assessment, 33*(3), 368-374. doi: 10.1007/s10862-010-9212-7

Wetter, M. W., Baer, R. A., Berry, D. T. R., Smith, G. T., Larsen, L. H. (1992). Sensitivity of MMPI-2 Validity Scales to random responding and malingering. *Psychological Assessment, 4,* 369-374.

Wolfe, R., & Hanley, J. (2002). If we're so different, why do we keep overlapping? When 1 plus 1 doesn't make 2. *Canadian Medical Association Journal, 166*(1), 65-66.

VITA

Joseph Brooks Minifie

Education

2010 – Present **Virginia Consortium Program in Clinical Psychology**
700 Park Avenue/MCAR-410
Norfolk, VA 23504
Psy.D., Clinical Psychology, expected **August 2015**
University-based, APA accredited program, jointly sponsored by: Eastern Virginia Medical School, Norfolk State University, and Old Dominion University

2013 **Norfolk State University**
Norfolk, VA 23504
M.A., Community/Clinical Psychology, **July 2013**

2008 **Trinity College**
Hartford, CT 06106
B.S., Neuroscience

Clinical Training

Predoctoral Internship:

Vanderbilt University-Department of Veterans Affairs Internship in Professional Psychology

Nashville, TN
Dates: July 2014 – June 2015

Postdoctoral Fellowship:

Clinical Psychology Postdoctoral Fellowship Program: Interprofessional Mental Health

Tennessee Valley Healthcare System
Alvin C. York VA Medical Center
Murfreesboro, TN
Dates: July 2015 – July 2016