

January 2016

# Comparability of SIRS and SIRS-2 Classifications with External Response Bias Criteria

Jessica Lynn Tylicki  
*Eastern Kentucky University*

Follow this and additional works at: <https://encompass.eku.edu/etd>

 Part of the [Psychology Commons](#)

---

## Recommended Citation

Tylicki, Jessica Lynn, "Comparability of SIRS and SIRS-2 Classifications with External Response Bias Criteria" (2016). *Online Theses and Dissertations*. 439.  
<https://encompass.eku.edu/etd/439>

This Open Access Thesis is brought to you for free and open access by the Student Scholarship at Encompass. It has been accepted for inclusion in Online Theses and Dissertations by an authorized administrator of Encompass. For more information, please contact [Linda.Sizemore@eku.edu](mailto:Linda.Sizemore@eku.edu).

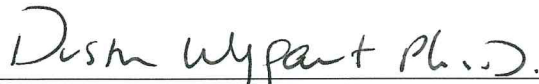
Comparability of SIRS and SIRS-2 Classifications with External Response Bias

Criteria

By

Jessica L. Tylicki

Thesis Approved:



---

Chair, Advisory Committee



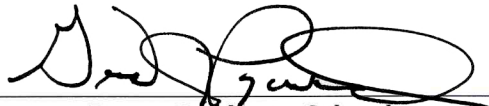
---

Member, Advisory Committee



---

Member, Advisory Committee



---

Dean, Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at Eastern Kentucky University, I agree that the Library shall make it available to borrowers under rules of the Library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of the source is made. Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor, or in his absence, by the Head of Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature Ashlea L. K. [Handwritten Signature]

Date 7/26/2016

Comparability of SIRS and SIRS-2 Classifications with External Response Bias  
Criteria

By

Jessica L. Tylicki

Bachelor of Science  
Eastern Kentucky University  
Richmond, Kentucky  
2016

Submitted to the Faculty of the Graduate School of  
Eastern Kentucky University  
in partial fulfillment of the requirements  
for the degree of  
MASTER OF SCIENCE  
August, 2016

Copyright © Jessica L. Tylicki, 2016  
All rights reserved

## DEDICATION

My thesis is dedicated to my parents, John and Sheree Tylicki, for providing me with a strong foundation only to excel in my various endeavors, and also to my brother, Matthew Tylicki, for always fulfilling each of my IT requests, especially from another state.

## ACKNOWLEDGMENTS

First and foremost, I express my greatest gratitude to Dr. Dustin Wygant for his valuable mentorship and for believing in me when I did not throughout the past two years. I thank my other committee members, Dr. Myra Beth Bundy and Dr. Michael McClellan, for their patience as I extended time to complete my thesis and for their worthy comments in my drafts. I must also thank my significant other, James Steriovski, for remaining by my side during both the ups and downs of this journey, even when you were over four-hundred miles away more so than not. Finally, I appreciate Dr. Robert Granacher, Dr. Richard De Mier, Dr. Elizabeth Tyner, Dr. Richard Frederick, and the Bureau of Prisons for allowing me to have access to their datasets.

## Abstract

The current study compared the classification groups between the SIRS and SIRS-2 using samples of disability claimants and criminal defendants. Results suggest that the newly revised SIRS-2 may have less clinical utility than the original SIRS. Implications of these results for both clinical and forensic settings are discussed.



## TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION.....	1
The Structured Interview of Reported Symptoms (SIRS) .....	4
Revisions of the Original SIRS .....	9
Prior Research.....	12
Concerns of the Revisions .....	15
Current Study.....	17
II. METHODS.....	19
Participants.....	19
Measures.....	19
III. RESULTS .....	21
IV. DISCUSSION.....	30
Limitations .....	32
Conclusion.....	33
References.....	35

LIST OF TABLES

TABLE		PAGE
1.	MMPI-2-RF Validity Scale Results for 37 Disability Claimants Classified as Feigning on the SIRS (Non-Content Based Invalid Results included).....	22
2.	MMPI-2-RF Validity Scale Results for 105 Criminal Defendants Classified as Feigning on the SIRS (Non-Content Based Invalid Results included) ...	23
3.	MMPI-2-RF Validity Scale Results for 36 Disability Claimants Classified as Feigning on the SIRS (Non-Content Based Invalid Results removed) .....	24
4.	MMPI-2-RF Validity Scale Results for 69 Criminal Defendants Classified as Feigning on the SIRS (Non-Content Based Invalid Results removed) .....	25
5.	% Above Cutoff for MMPI-2-RF Validity Scale Results for 37 Disability Claimants Classified as Feigning on the SIRS (Non-Content Based Invalid Results included).....	26
6.	% Above Cutoff for MMPI-2-RF Validity Scale Results for 105 Criminal Defendants Classified as Feigning on the SIRS (Non-Content Based Invalid Results included).....	27
7.	% Above Cutoff for MMPI-2-RF Validity Scale Results for 36 Disability Claimants Classified as Feigning on the SIRS (Non-Content Based Invalid Results removed) .....	28
8.	% Above Cutoff for MMPI-2-RF Validity Scale Results for 69 Criminal Defendants Classified as Feigning on the SIRS (Non-Content Based Invalid Results removed) .....	29

## CHAPTER I

### INTRODUCTION

Psychological assessments can provide valuable data to various professionals seeking specific information about an individual. Unfortunately, those undergoing psychological evaluations can compromise test results by over-reporting and even blatantly misrepresenting physical or psychological symptoms, which leaves evaluating professionals with an inaccurate clinical impression. The tendency for individuals to report false symptoms is typically motivated by external incentives, such as evasion of criminal responsibility (e.g., insanity defense) or financial gains (e.g., disability compensation), which is particularly problematic in forensic settings. Notably, about one in six examinees evaluated in a forensic context is suspected of malingering symptoms (Rogers, Salekin, Sewell, Goldstein, & Leonard, 1998). In another widely cited study, base rates of malingering and symptom exaggeration were about 29%, 30%, 19%, and 8% in personal injury, disability, criminal, and psychiatric cases, respectively (Mittenberg, Patton, Canyock, & Condit, 2002). Inaccurate reporting can significantly impact test results, which, in turn, can impact the course of treatment and legal decisions. Protocol validity, which refers to the accuracy of an individual's test results as it pertains to their psychological functioning, is assessed through the use of validity scales and is significantly affected by an over-reporting of symptoms. For instance, if an individual intentionally endorses an unusual amount of symptoms on the Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008/2011), this inaccurate approach of responding can inflate the Restructured Clinical (RC)

scales. Therefore, RC scale scores would overestimate psychological problems and symptoms. Accurate interpretation of these validity scales, which may indicate unsuccessful malingering of symptoms, can lead to penalties, ranging from termination of mental health treatment to a reduction in benefits and negative outcomes in litigation.

There are a variety of factors that may weaken the validity and utility of psychological assessments. The usefulness of psychological test data can be compromised by specific response styles that invalidate personality and psychopathology test results. Content-based invalid responding is one threat to protocol validity that occurs when examinees pay attention to the content of the items and provide responses that portray a distorted picture of their functioning (Burchett & Bagby, 2014). This type of responding may be intentional or unintentional and includes *over-reporting*, which occurs when examinees attempt to appear worse than they actually are. Unintentional over-reporting may be due to poor insight into an examinee's symptoms or negative emotionality, which predisposes individuals to believe they are more impaired than they actually are (Ben-Porath, 2013). *Feigning* is an intentional exaggeration or fabrication of symptoms, but no assumptions are made about an examinee's motivation to do so, while *dissimulation* refers to intentional response distortion (Burchett & Bagby, 2014). *Malingering*, which is a sub-category of feigning, is defined in the DSM-5 (APA, 2013) as the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external, secondary gain. Wygant and Granacher (2015) explained how individuals who malingering may display or report

symptoms that are not accurate representations of specific disorders. For example, an individual malingering symptoms of Schizophrenia in an effort to “play crazy” might only focus on hallucinations and delusions, which are obvious symptoms and popular notions of psychosis, while overlooking negative symptoms of the disorder (e.g., anhedonia). Thus, the approach in which individuals malingering psychopathology is often due to a lack of understanding of the unique characteristics of a disorder. In summary, content-based invalid responding may occur in three symptom domains, including somatic complaints, cognitive complaints, and psychopathology (Rogers et al., 2010). Because different types of response distortion influence the accuracy of reported symptoms, it is critical for professionals to examine validity scales and utilize other assessments that can detect the possibility of inaccurate reporting of symptoms.

The Structured Interview of Reported Symptoms (SIRS; Rogers, Bagby, & Dickens, 1992) is one of the most widely used (Archer, Buffington-Vollum, Stredny, & Handel, 2006) and empirically validated measures of feigned psychopathology (Green & Rosenfeld, 2011). In 2010, the SIRS-2 (SIRS, 2<sup>nd</sup> edition; Rogers, Sewell, & Gillard, 2010), which included significant revisions, was released. However, there have been several criticisms about the revised version that concern its clinical utility and whether or not the SIRS-2 should replace the SIRS. Therefore, it is important that independent research is conducted in an effort to address these concerns. This thesis project will examine the utility of the SIRS-2 classification scheme in relation to the original SIRS, as well as in relation to external markers of symptom

exaggeration. This project will examine data from two separate forensic settings, one involving civil litigants and one involving criminal defendants.

### **The Structured Interview of Reported Symptoms (SIRS)**

The SIRS (Rogers et al., 1992) is a structured interview intended to assess whether an individual is feigning symptoms of psychopathology. The original SIRS included 172 items that were developed to provide data on feigning and honest responding, in addition to insights into how a client feigns symptoms in cases of dissimulation (DeClue, 2011). Items are designated as “Detailed Inquires,” which include questions about specific psychological problems and are repeated later. If examinees rate the same question with a different response, then they are given one point that totals into the *Inconsistency of Symptoms* Supplementary Scale. Other items are designated as “General Inquires,” which include items about broad psychological problems. Most items yield a score of 0, 1, or 2, while the remaining items, which assess problems of concentration and memory, yield total errors. There are eight primary scales on the SIRS and five supplementary scales. The primary scales represent detection strategies that are organized into two general categories, including unlikely symptoms or amplified symptoms. (Rogers, Payne, Berry, & Granacher, 2009).

The Primary Scales from “General Inquiries” include unlikely symptoms. These reflect a presentation that is very atypical of mental illness and endorsement of these items is indicative of feigning. They include the following scales as described by Rogers, Gillis, Dickens, and Bagby (1991):

Rare Symptoms (RS): Items that comprise this scale assess for symptoms that are infrequently observed in psychiatric patients.

Symptom Combination (SC): Items in this scale ask about the presence of genuine psychiatric problems that rarely occur simultaneously. The exclusive presence of a symptom may be frequent but the combination of symptoms is infrequent.

Improbable or Absurd Symptoms (IA): The outrageous quality of these symptoms distinguished them from rare symptoms and makes it highly unlikely that these symptoms are valid.

Reported vs. Observed Symptoms (RO): Items of this scale ask about behaviors, specifically physical movements and speech that are compared with clinical observations.

The remaining four Primary Scales from “Detailed Inquiries” include amplified symptoms, which reflect a presentation of more realistic problems, but a suspicion of feigning arises when items are endorsed at a higher frequency or intensity than is typical of mental illness (Rogers et al., 2009). They include the following scales as described by Rogers and colleagues (1991):

Blatant Symptoms (BL): These symptoms, which are typically over-endorsed by individuals malingering psychopathology, are obvious signs of a mental disorder. Blatant symptoms are those that naïve individuals would likely identify as indicative of a mental disorder.

Subtle Symptoms (SU): These are symptoms that naïve individuals would likely identify as daily problems and not indicative of mental illness.

Severity of Symptoms (SEV): Given that individuals malingering psychopathology

often endorse an unlikely number of symptoms with extreme or unbearable severity, examinees are asked, from a subset of 32 symptoms, which are “unbearable” or “too painful to stand.”

Selectivity of Symptoms (SEL): Some individuals malingering psychopathology are nonselective or indiscriminant in their endorsement of psychiatric problems. This scale is an overall measurement of symptom endorsement from a subset of 32 symptoms.

Supplementary scales serve to provide clinical descriptions beyond feigning by addressing response consistency and defensiveness, in addition to a willingness to acknowledge poor relationships with mental health professionals and a tendency to exaggerate or fabricate symptoms (Rogers et al., 1992). They include the following scales:

Direct Appraisal of Honesty (DA): A high score suggests that the examinee acknowledges a lack of openness with mental health professionals and a tendency to exaggerate psychological problems.

Defensive Symptoms (DS): Items of this scale include various daily problems, worries, and situations, which most individuals experience to some degree, and denial of these symptoms may be an indicator of defensiveness.

Improbable Failure (IF): This scale is intended as a screener for feigned cognitive impairment. It is designed as a set of simple cognitive tasks, such as opposites and rhyming, that can be easily completed by those who speak English as their primary language and who have been exposed to these tasks in their primary education.



Overly Specified Symptoms (OS): This scale assesses the endorsement of symptoms with an unrealistic degree of precision. A high score raises concerns about the accuracy of the examinee's reporting.

Inconsistency of Symptoms (INC): This scale is a measure of the consistency of the examinee's self-report. However, inconsistencies should not necessarily be considered as evidence of feigning.

Following administration, a specific scoring and classification method is utilized. Examinees are classified as feigning when they score in the *Definite* feigning range on one primary scale or in the *Probable* feigning range on three primary scales. For Indeterminate cases, which occurs when an examinee scores in the *Probable* feigning range on one or two scales and/or in the *Indeterminate* range on several scales, the Total Score (greater than or equal to a raw score of 76) can be used to classify feigning.

Previous research has examined psychometric properties of the SIRS. Rogers and colleagues (1991) examined the usefulness of the SIRS in identifying dissimulators when compared to psychiatric outpatients and inpatients. First, they found a high level of discrimination for the SIRS scales between simulators and community and outpatient control participants. Results also provided support for concurrent validity of the SIRS scales confirmed by indices of malingering and item consistency on the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1943). To determine whether individuals suspected of malingering psychopathology responded differently to the SIRS than non-malingering psychiatric inpatients, the authors found that those suspected of malingering

psychopathology endorsed significantly more symptoms than non-malingering inpatients. However, the two groups did not differ in the extent that they admitted to everyday problems (DS) and less-than-honest interactions with others (DA), or in the frequency with which they endorsed overly specified symptoms (OS) and atypical symptom onset (SO). As a result, some SIRS scales were not equally effective at differentiating feigners from clinical subjects. Despite these results, Rogers and colleagues (1991) still maintain that the SIRS appears to be a reliable and valid measure in the assessment of malingering, since the SIRS is useful in identifying specific dissimulation styles found in psychiatric patients.

In a later study, SIRS scores in the feigning range were most strongly associated ( $r = .81$ ) with the Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997) total score, which serves as an indicator of feigned psychopathology (Edens, Poythress, & Watkins-Clay, 2007). The SIRS scores correlated slightly less ( $r = .75$ ) with high cutoff scores on the Negative Impression Management scale (NIM) of the Personality Assessment Inventory (PAI; Morey, 1991), which includes items that represent an exaggerated impression or extremely bizarre and unlikely symptoms. Edens and colleagues also found smaller correlations with similar indicators of malingering of the PAI, including the Malingering Index (MAL; Morey, 1996) ( $r = .57$ ) and the Rogers Discriminant Function (RDF; Rogers, Sewell, Morey, & Ustad, 1996) ( $r = .40$ ).

Despite the high intercorrelations among some of the scales, rates of sensitivity and specificity within the four subsamples varied significantly. For participants from a prison's mental health unit, sensitivity was .50 for suspected

individuals malingering psychopathology, who were judged on admission to be feigning symptoms of psychopathology., Specificity was .60 for patients, who were judged by psychiatrists to be genuinely mentally ill (Edens et al., 2007). However, sensitivity for simulators, who were instructed to feign serious mental illness was .90 and specificity for controls, who completed measures under standard instructions, was 1.0. These results suggest that cutoff scores for validity indicators were much less accurate in classifying mental health unit participants than general population participants. Furthermore, each index correctly identified a greater percentage of general population inmates instructed to malingering than mental health unit inmates judged clinically to be feigning. Overall, the SIRS yielded a poor sensitivity rate in the forensic sample.

### **Revisions of the Original SIRS**

Rogers, Sewell, and Gillard (2010) released a revised version of the SIRS, the SIRS-2 (SIRS, 2<sup>nd</sup> edition), which introduced significant modifications to the classification decision model for interpreting performance on the SIRS-2. These modifications were essentially aimed at reducing false positives. The SIRS-2 remains a highly structured interview and items have not been changed. The SIRS-2 manual includes updated normative data for the primary scales, which used original SIRS validation research that was conducted prior to 1992, and also research conducted after the publication of the SIRS. However, validation data for the new scoring guidelines of the SIRS-2 were derived from only 522 cases, such that 314 cases were drawn from the original validation studies and 208 cases were drawn from a sample of multiply traumatized inpatients (Green et al., 2012). Similar to the original SIRS,

the SIRS-2 consists of primary scales, supplementary scales, and a classification scale. The eight primary scales are used for both clinical descriptions and the classification of response styles, while the Supplementary Scales are used mostly for a descriptive analysis of scale elevations (Rogers et al., 2010). The SIRS-2 includes the same eight Primary Scales and five Supplementary Scales. However, while the Improbable Failure (IF) scale maintains the same 20 items as the original SIRS, the SIRS-2 divides them into four items with five components each (Rogers et al., 2010).

More importantly, the SIRS-2 added a new classification scale, the *Rare Symptom-Total* (RS-Total) scale, which was constructed to differentiate between feigned or atypical and genuine clinical presentations (Rogers et al., 2010) and is used to minimize false positives. Items that are not keyed on the Primary Scales and have low rates of endorsements among presumed genuine clinical patients are included in the RS-Total scale. Furthermore, two new indices, the *Modified Total Index* (MT Index or MTI) and *Supplementary Scale Index* (SS Index), were added to the SIRS-2. The MTI replaced the original SIRS Total score and is used to determine a Feigning or Indeterminate classification. It is the sum of four Primary Scales, including RS, SC, IA, and BL, given their large effect sizes in discriminating genuine and feigning protocols with few false positive errors (Rogers et al., 2010). The SS index was constructed to assess a strategy of disengagement, where an examinee may attempt to disrupt the interview by providing non-affirmative responses (Kocsis, 2011). It is the sum of four Supplementary Scales, including DA, DS, IF, and OS.

The decision model, which is used to assist clinicians in determining appropriate classifications of the examinee, is another significant revision of the original SIRS. Interpreting the interview data requires following a procedure of steps that are displayed on the inside of SIRS-2 booklets. The first step of interpretation requires establishing whether a certain number of Primary Scales fall in the definite or probable ranges. More specifically, if three or more Primary Scales fall in the probable range or one or more falls in the definite range, then an examinee is classified as Feigning. If neither of these criteria are satisfied, then the next step involves determining whether at least one Primary Scale falls in the probable range. If so, and the examinee's MTI is above 45, then the examinee is classified as Feigning. If the MTI is below 45, then other sub-steps are followed to determine if the examinee is classified as Indeterminate, Disengagement, or Genuine Responding by further evaluating the MTI or SS Index. However, if at least one Primary Scale does not fall in the probable range and the SS Index is below four, the examinee is classified as Disengagement; if the SS Index is not below four, the examinee's response style is classified as Genuine Responding. The five classifications are further described in detail below by Rogers and colleagues (2010):

Genuine: the examinee has scores that are strongly characteristic of an individual who made no effort to over-report symptoms.

Feigning: the examinee displays a pattern of scores that are characteristic of an individual feigning a mental disorder. The examinee must have elevations on primary scales and on either the RS-Total scale or the MT Index.

Indeterminate-Evaluate: the examinee has at least one primary score elevated but the assessment was inconclusive. Since the likelihood of feigning in this category exceeds 50%, further examination is required. The examinee must score in the probable range of one or two primary scales and have an MT Index ranging from 22-45.

Indeterminate-General: the examinee has at least one primary score elevated, but does not indicate an increased likelihood of feigning. The examinee must score in the probable range of one or two primary scales and have an MT Index ranging from 13-21.

Disengagement Response Style: the examinee was minimally involved throughout the administration. This style is typically used by feigners to avoid the detection of the feigned mental disorders. Since the likelihood of feigning in this category exceeds 50%, further examination is required. The examinee must have a score below four on the SS Index.

### **Prior Research**

There have been several studies that have evaluated the classification accuracy of the SIRS-2. Green, Rosenfeld, and Belfi (2012) directly compared classifications made by the SIRS-2 with the original SIRS. They used a criterion group, which included archival records of 114 pretrial criminal defendants admitted to a forensic psychiatric center for restoration of competency to stand trial. Participants within this sample were primarily males and ethnic minorities with long histories of mental illness. The simulation group included 36 community members with similar demographic and background characteristics to the forensic

sample in terms of gender, ethnicity, and socioeconomic status. However, participants within the community sample had significantly more years of education, fewer arrests, and were less likely to have been diagnosed with a mental disorder. Those admitted to the forensic hospital completed a standardized battery of measures to determine whether they were competent to stand trial and therefore, be discharged. Furthermore, prior to testing, treating psychiatrists informed evaluators whether or not they believed participants were feigning any symptoms. The community sample was administered the same battery of measures as the forensic sample. They were also offered an incentive to simulate an incompetence to stand trial presentation by persuasively feigning symptoms of mental illness while evading detection from the measures.

For the forensic psychiatric participants, more were classified as Indeterminate or Feigning by the SIRS than the SIRS-2. Similarly, community simulators were more likely to be classified as feigning by the SIRS than the SIRS-2, even though this was not a significant finding. Fewer simulators fell in the Indeterminate categories with the SIRS than the SIRS-2. When examining the new scoring method of the SIRS-2, two of seven genuine patients with elevated primary scales were eliminated from being classified as Feigning. However, two patients suspected of feigning and three simulators, who exceeded the threshold of primary scale elevations, did not obtain scores above the cutoff on the RS-Total scale and therefore, were classified as either Indeterminate or Genuine. These findings suggest that the new RS Total reduced false positive classifications, but at the cost of reduced sensitivity. As a result, it must be considered whether or not the optimal

cut-offs for classifications should be reduced. Additionally, scores on the MT-Index were reviewed for those cases with unclear evidence of feigning, and results showed that the MT-Index did not categorize any participants as likely feigning. Thus, it is questionable whether the MT-Index is actually useful in differentiating feigners from genuine responders (Green et al., 2012).

While there are many strengths of the above study, the results may be limited by the study design, such that the use of criterion groups to classify individuals as feigning or genuine may introduce an unknown amount of error, such as inaccurate calculations in true positive and false positive rates. However, the authors mention a few steps that were taken to minimize the possibilities of error. They required a consensus between psychiatrists' classifications and evaluators' classifications in the forensic sample. Limitations of the study design were also addressed by including a community sample that was comparable to the forensic sample. Of course, these precautions did not eliminate all sources of errors.

A more recent study examined how often elevations on the M-FAST (Miller-Forensic Assessment of Symptoms Test; Miller, 2001) associated with elevations on the SIRS and SIRS-2 (Glassmire, Tarescavage, & Gottfried, 2016). They found that for individuals who elevated on the M-FAST over the recommended cutoff, 66.0% met SIRS criteria for feigning, while only 42% met SIRS-2 criteria for feigning. The discordance between the M-FAST and SIRS-2 with the SIRS further highlights the reduction in sensitivity of the SIRS-2. This finding serves as a critical issue for clinicians, such that they are likely to observe more discrepant test findings between the M-FAST and SIRS-2 compared to the SIRS.



## Concerns of the Revisions

The developers of the SIRS-2 recommend that it be used rather than the original SIRS, so it is expected that the SIRS-2 would replace the SIRS as the premier measure of feigned mental disorders (DeClue, 2011). It should be noted that more recently, the original SIRS is no longer commercially available by its distributor, Psychological Assessment Resources, Inc. However, since the revisions of the original SIRS, there have been several additional concerns about the effectiveness of the SIRS-2 for its primary clinical purpose.

Tyner and Frederick (2011) argued that scoring methods for the SIRS and SIRS-2 produce computational errors that affect the overall classification of primary scale scores. They found that 17 records had at least one misclassification, and after the primary scales were correctly classified, only one changed an overall classification. Thus, it is clear that even a miscount of one point for a scale can change an overall classification. The susceptibility of examiners to make errors further impacts the classification accuracy of the SIRS-2.

Green and Rosenfeld (2011) noted that the SIRS-2 represents an effort to improve specificity by applying strict criteria for identifying Feigning and ambiguous cases as Indeterminate. However, the SIRS-2 manual has not yet been subjected to cross-validation. As a result, the extent to which the classification accuracy has been improved is still unknown.

DeClue (2011) noted that the authors of the SIRS-2 (Rogers et al., 2010) failed to identify the subjects involved in the validation of the SIRS-2 data set. They do not provide any demographics about the subjects and whether there were

differences in characteristics between the subjects. Furthermore, Rubenzer (2010) stated that Rogers advocates for known-groups designs, where the criterion groups (feigning vs. honest patients) are accurately defined. However, the criterion by which the SIRS-2 was validated is minimally described in the test's manual. It seems to be based on the clinical judgment of a clinician, and there is no clear description of what information or processes aided in these decisions (Rubenzer, 2010). Information about validation can influence the outcome of a study, but other researchers cannot independently examine it as it remains unknown.

Moreover, Rubenzer (2010) stated that the authors did not include subjects who could not be reliably classified on the SIRS-2, and the number of Indeterminate cases was substantial. It was questioned whether sensitivity and specificity are appropriately calculated when there are Indeterminate cases. With such cases, the estimate of diagnostic statistics, particularly sensitivity, would be unknown (Rubenzer, 2010). Furthermore, the statistics reported by the authors were calculated incorrectly. The manual reported the sensitivity to be .80 with a false-positive rate of 2.5%, but after re-calculation, Rubenzer (2010) stated that the false-positive rate is actually 3.6%.

Additionally, while the Improbable Failure (IF) scale maintains the same items from the SIRS, the authors of the SIRS-2 (Rogers et al., 2012) imply extensive validation of the scale, but no references were provided (Rubenzer, 2010). Rogers and colleagues (2010) also noted that the scale's utility may be limited to those who do not have impaired intellectual functioning. As a result, examinees with an IQ below 80 were significantly more likely to make errors than those with an IQ above

80 (Rubenzer, 2010). Moreover, non-psychotic disorders (e.g., anxiety, depression, PTSD) were reported to have limited effects on scores of the IF scale, but the effect of psychosis, use of drugs, or traumatic brain damage was not discussed either. Therefore, the utility of the IF scale may be restricted to those without intellectual disabilities or neurocognitive disorder. As discussed earlier, the SS Index contains four of the SIRS-2 Supplementary Scales, but there is no previous data to support the composition of the SS Index or report on its internal consistency or incremental validity compared to the Defensive Symptoms scales (Rubenzer, 2010). Further research is needed to examine the validity and effectiveness of the SS Index.

Overall, several significant concerns have been presented about the effectiveness of the SIRS-2. The nature of its scoring method influences the classification accuracy of the SIRS-2. There is no information available to allow others to validate the authors' findings; therefore, the SIRS-2 classification rules have not been cross-validated. Furthermore, the manual reports a sensitivity of the SIRS-2 as if there were no Indeterminate cases. The false-positive rate and specificity reported in the manual appear to be incorrect. Lastly, the IF scale may be limited in its use and there is no data to support the SS Index. These concerns suggest that the SIRS-2 revision may have less clinical utility than the original SIRS.

### **Current Study**

To address the above concerns, independent research is warranted to cross-validate data regarding the classification accuracy of the SIRS-2. The current study examined the classification rates of the SIRS-2 in relation to the original SIRS. To my knowledge, this is the first study to examine SIRS and SIRS-2 feigning rates in civil

and criminal forensic samples using external response bias criteria, namely the MMPI-2-RF validity scales. This study will enhance the prior study of Green and colleagues (2012), as we included two unique samples of participants, who were motivated by real-life external gains to feign symptoms. In contrast, the previous study included a simulation group of community members, who were motivated by a monetary incentive to feign symptoms of mental illness that may not have been an accurate presentation otherwise observed in a forensic setting. The civil and criminal samples within this study were also evaluated in regards to different referral questions, so each sample may have portrayed more or less feigned symptoms of psychopathology and physical symptoms as distinct presentations. This should have allowed our study to capture and analyze a broader range of possible portrayals of feigning rather than focusing on participants from one setting.

For my hypotheses, I predicted that in both samples, fewer participants would be classified as Feigning on the SIRS-2 than the SIRS and more participants would be reclassified as Indeterminate on the SIRS-2 than the SIRS. Given that Green and colleagues (2012) found that the SIRS-2 yielded lower sensitivity rates, I further predicted that for those who are reclassified in either Indeterminate categories and in the Honest category from the Feigning category of the original SIRS, MMPI-2-RF validity scales would suggest that these participants are in fact, responding in an inconsistent manner and over-reporting symptoms. In other words, while the SIRS-2 may indicate that participants may not be feigning, MMPI-2-RF validity scales will suggest otherwise. I predicted that this pattern of results will be similar across both samples.

## Chapter II

### METHODS

#### **Participants**

The civil sample includes 251 males (65.2%) and 134 females (34.8%) for a total of 385 litigants, who were referred for psychological evaluations in Lexington, Kentucky. The mean age of participants was 41.5 years ( $SD = 10.8$ ) with a mean education of 12.16 years ( $SD = 2.0$ ). Participants were primarily Caucasian (95.4%), and 4.6% were African-American. Referral questions included worker's compensation and disability determinations, and all of the participants underwent a complete psychological evaluation as part of their litigation.

Archival data from 140 competency to stand trial/criminal responsibility referrals were used within the criminal sample. All participants were evaluated at the Federal Medical Center in Springfield, Missouri after being charged with various types of felony Federal offenses. All participants were male with a mean age of 34.10 ( $SD = 8.4$ ). The sample was primarily African American followed by Caucasian and other ethnicities. They completed a battery of tests, but the current study focused on data from a select few as described below.

#### **Measures**

The current study analyzed data from the same measures in both the criminal and civil forensic samples. The measures are described below:

*Structured Interview of Reported Symptoms.* The SIRS (Rogers et al., 1992) is a structured interview with 172 items designed to assess feigned mental disorders using specific detection strategies as discussed in earlier paragraphs. Participants

were administered the original SIRS from which the SIRS-2 primary, secondary, and classification scales were calculated. For the criminal sample, participants were only administered the SIRS if they were suspected of feigning by unit staff.

*Minnesota Multiphasic Personality Inventory-2 Restructured Form.*

Participants were initially administered the MMPI-2 (Butcher et al., 2001) from which the MMPI-2-RF scales were scored. The MMPI-2-RF (Ben-Porath & Tellegen, 2008/2011) includes 338 self-report items consisting of true-false items. It was designed to assess adult personality and psychopathology. The current study only utilized the validity scales of the MMPI-2-RF.

## Chapter III

### RESULTS

For those who were originally classified as Feigning on the SIRS, the first step of analyses evaluated the mean scores of MMPI-2-RF validity scales for participants within each SIRS-2 classification. SIRS-2 classifications were calculated using item responses from each participant's SIRS protocol. Table 1 and Table 2 included non-content based invalid results from the MMPI-2-RF for the civil and criminal samples, respectively. The current study evaluated results with and without non-content based invalid results to determine how many individuals within each SIRS-2 classification continued to display evidence of over-reporting, particularly when their invalid MMPI-2-RF results were removed. In the civil forensic sample, 37 participants were originally classified as Feigning on the SIRS, but only 5 participants remained within the Feigning classification on the SIRS-2. Half of the remaining participants ( $n = 16$ ) were classified as Honest, while the other 50% of participants were classified into the Indeterminate categories. Notably, mean scores on the MMPI-2-RF validity scales displayed strong evidence of over-reporting across all related scales for those in each of the four classifications, which is presented in Table 1. For example, each classification group yielded a mean F-r score above 100T, which suggests that their MMPI-2-RF protocols may be invalid due to over-reporting of psychological dysfunction. Those within the Honest and Indeterminate-General classification groups yielded a mean RBS score above 100T, which suggests the possibility of response bias with respect to self-reported cognitive complaints.

Table 1.

MMPI-2-RF Validity Scale Results for 37 Disability Claimants Classified as Feigning on the SIRS (Non-Content Based Invalid Results included)

	VRIN-r	TRIN-r	F-r	Fp-r	Fs	FBS-r	RBS
SIRS-2 Classification							
Feigning (n = 5)	50.20 (10.6)	59.00 (10.0)	118.20 (4.0)	76.60 (8.5)	102.20 (13.4)	95.00 (6.2)	111.20 (9.7)
Genuine (n = 16)	48.94 (9.4)	61.38 (11.1)	109.00 (11.6)	70.69 (22.1)	89.63 (22.6)	86.75 (16.8)	100.94 (16.3)
Ind-Evaluate (n = 2)	58.00 (7.1)	57.00 (0.0)	110.50 (13.4)	63.50 (6.4)	82.50 (34.6)	97.50 (2.1)	92.50 (23.3)
Ind-General (n = 14)	47.43 (7.1)	60.29 (8.8)	113.36 (10.4)	81.43 (25.7)	91.14 (23.8)	91.00 (10.3)	109.93 (9.8)

*Note:* SD in parentheses below the mean scores. *VRIN-r* variable response inconsistency, *TRIN-r* true response inconsistency, *F-r* infrequent responses, *Fp-r* infrequent psychopathology responses, *Fs* infrequent somatic responses, *FBS-r* symptom validity, *RBS* response bias scale, *L-r* uncommon virtues, *K-r* adjustment validity.

As shown in Table 2, 105 criminal defendants were initially classified as Feigning on the SIRS, whereas approximately 50% (n = 52) remained classified as Feigning on the SIRS-2. Mean scores on the MMPI-2-RF validity scales revealed evidence of infrequent responding (F-r) for symptoms of psychopathology (Fp-r) and somatic symptoms (Fs), and over-reporting of cognitive and memory complaints (FBS-r, RBS). Participants who were reclassified as Indeterminate-Evaluate (n = 34) had slightly higher MMPI-2-RF mean scores than those reclassified as Indeterminate-General (n = 15), while both groups had lower mean MMPI-2-RF validity scores compared to those in the Feigning classification. Participants reclassified as Honest (n = 4) had the lowest mean MMPI-2-RF validity scores with the exception of the infrequent responding (F-r) scale, which had the highest mean score of all four classification groups.



Table 2.  
MMPI-2-RF Validity Scale Results for 105 Criminal Defendants Classified as Feigning on the SIRS (Non-Content Based Invalid Results included)

SIRS-2 Classification	VRIN-r	TRIN-r	F-r	Fp-r	Fs	FBS-r	RBS
Feigning (n = 52)	64.87 (17.7)	66.15 (14.2)	116.15 (9.7)	111.02 (14.4)	97.98 (21.7)	84.42 (14.8)	103.98 (14.4)
Genuine (n = 4)	50.75 (8.2)	63.25 (12.7)	118.75 (2.5)	98.25 (17.9)	64.00 (15.1)	65.50 (10.1)	86.00 (16.3)
Ind-Evaluate (n = 34)	61.53 (15.7)	61.47 (11.5)	115.26 (13.1)	109.35 (15.6)	97.21 (19.2)	82.82 (16.4)	103.03 (16.4)
Ind-General (n = 15)	62.40 (18.6)	67.13 (19.7)	113.87 (12.1)	102.93 (15.9)	87.73 (23.3)	75.33 (10.84)	98.60 (16.2)

*Note:* SD in parentheses below the mean scores. *VRIN-r* variable response inconsistency, *TRIN-r* true response inconsistency, *F-r* infrequent responses, *Fp-r* infrequent psychopathology responses, *Fs* infrequent somatic responses, *FBS-r* symptom validity, *RBS* response bias scale, *L-r* uncommon virtues, *K-r* adjustment validity.

During the next set of analyses, non-content based invalid results were removed based on cut-off scores (*VRIN-r* and *TRIN-r* > 80T), and mean scores were re-calculated for MMPI-2-RF validity scales in both the civil and criminal samples. As shown in Table 3, only one participant from the civil sample was excluded after invalid protocols were removed. While 36 participants were classified as Feigning on the original SIRS, only one individual remained classified as Feigning on the SIRS-2. As shown in Table 4, 36 criminal defendants were excluded. Of the 69 participants who remained classified as Feigning on the SIRS, about 46% (n = 32) stayed within this classification group on the SIRS-2. For both samples, mean MMPI-2-RF validity scores continued to display evidence of infrequent responding and over-reporting of various symptoms across all four classification groups. For example, those within the Honest classification groups in both samples yielded similar mean *F-r* scores as those in the same groups before invalid results were removed. For both samples,

participants within both of the Indeterminate groups yielded the highest mean scores on the F-r scale compared to all other scales. Their mean scores on this scale suggest that their protocols may be invalid due to over-reporting of psychological dysfunction. Interestingly, participants within the criminal sample displayed significantly higher mean scores on the Fp-r scale compared to those in the civil sample. Furthermore, about 42% (n = 15) of participants within the civil sample and about 4% (n = 3) of participants within the criminal sample were reclassified as Honest on the SIRS-2 after originally being classified as Feigning on the SIRS.

Table 3.  
MMPI-2-RF Validity Scale Results for 36 Disability Claimants Classified as Feigning on the SIRS (Non-Content Based Invalid Results removed)

SIRS-2 Classification	VRIN-r	TRIN-r	F-r	Fp-r	Fs	FBS-r	RBS
Feigning (n = 1)	63.00	73.00	111.00	68.00	115.00	86.00	105.00
Genuine (n = 15)	47.70 (5.9)	59.60 (8.9)	109.80 (11.5)	67.40 (18.4)	87.60 (21.9)	88.27 (16.2)	102.33 (15.8)
Ind-Evaluate (n = 6)	50.67 (9.5)	56.00 (5.6)	116.83 (7.8)	73.67 (10.5)	93.50 (20.4)	97.33 (3.4)	106.00 (16.9)
Ind-General (n = 14)	47.43 (7.1)	60.29 (8.8)	113.36 (10.4)	81.43 (25.7)	91.14 (23.8)	91.00 (10.3)	109.93 (9.8)

Note: SD in parentheses below the mean scores. *VRIN-r* variable response inconsistency, *TRIN-r* true response inconsistency, *F-r* infrequent responses, *Fp-r* infrequent psychopathology responses, *Fs* infrequent somatic responses, *FBS-r* symptom validity, *RBS* response bias scale, *L-r* uncommon virtues, *K-r* adjustment validity.

To provide another perspective of the results, the next set of analyses assessed how many participants who were classified as Feigning on the original SIRS scored above a chosen cutoff on each MMPI-2-RF validity scale after being reclassified on the SIRS-2. Table 5 and Table 6 included non-content based invalid results from the MMPI-2-RF for the civil sample and criminal forensic samples, respectively. For the civil sample, 80% (n = 4) of participants within the Feigning

Table 4.

MMPI-2-RF Validity Scale Results for 69 Criminal Defendants Classified as Feigning on the SIRS (Non-Content Based Invalid Results removed)

	VRIN-r	TRIN-r	F-r	Fp-r	Fs	FBS-r	RBS
SIRS-2 Classification							
Feigning (n = 32)	56.63 (11.0)	58.63 (8.0)	116.66 (9.5)	110.66 (13.6)	99.56 (21.0)	86.66 (15.2)	106.44 (12.8)
Genuine (n = 3)	50.00 (9.8)	57.67 (7.5)	118.33 (2.9)	91.00 (12.8)	60.67 (16.7)	64.00 (11.8)	79.67 (12.5)
Ind-Evaluate (n = 25)	56.64 (11.7)	57.80 (7.2)	114.12 (15.0)	109.76 (15.1)	99.16 (17.6)	82.00 (16.9)	102.32 (18.3)
Ind-General (n = 9)	56.33 (11.3)	55.78 (7.4)	109.78 (14.4)	96.56 (16.7)	76.11 (21.5)	71.33 (9.5)	94.67 (16.5)

*Note:* SD in parentheses below the mean validity scores. *VRIN-r* variable response inconsistency, *TRIN-r* true response inconsistency, *F-r* infrequent responses, *Fp-r* infrequent psychopathology responses, *Fs* infrequent somatic responses, *FBS-r* symptom validity, *RBS* response bias scale, *L-r* uncommon virtues, *K-r* adjustment validity.

classification group produced invalid protocols as they yielded a score of 120 on the F-r scale. More importantly, 64.3% (n = 9) of participants within the Indeterminate-General and 50% (n = 2) of participants within the Indeterminate-Evaluate groups also produced invalid protocols despite their indeterminate classifications on the SIRS-2. Additionally, 31.3% (n = 5) of those within the Honest group also produced invalid MMPI-2-RF protocols due to over-reporting. In table 6, for participants who were classified as Feigning on the SIRS-2, 82.7% (n = 43) and 80.8% (n = 42) produced invalid MMPI-2-RF protocols as they yielded a score of 120 on the F-r scale and a score of 100 on the Fp-r scale, respectively. Furthermore, 85.3% (n = 29) of those reclassified as Indeterminate-Evaluate and 73.3% (n = 11) of participants reclassified as Indeterminate-General also produced invalid protocols, given they yielded a score of 120 of the F-r scale. Notably, 75% (n = 3) of participants also produced invalid protocols due to infrequent responding (F-r), even though they

were reclassified as Honest on the SIRS-2. 50% (n = 2) of individuals within this same classification group produced invalid protocols due to over-reporting symptoms of psychopathology (Fp-r).

Table 5.  
% Above Cutoff for MMPI-2-RF Validity Scale Results for 37 Disability Claimants Classified as Feigning on the SIRS (Non-Content Based Invalid Results included)

	VRIN-r	TRIN-r	F-r	F-r	Fp-r	Fp-r	Fs	Fs	FBS-r	FBS-r	RBS	RBS
MMPI-2-RF Cutoff	80	80	100	120	90	100	90	100	90	100	90	100
SIRS-2 Classification												
Feigning (n = 5)	0 (0)	0 (0)	100 (5)	80 (4)	0 (0)	0 (0)	80 (4)	40 (2)	80 (4)	20 (1)	100 (5)	80 (4)
Genuine (n = 16)	0 (0)	6.3 (1)	81.3 (13)	31.3 (5)	25.0 (4)	31.3 (5)	43.8 (7)	31.3 (5)	37.5 (6)	25.0 (4)	81.3 (13)	62.5 (10)
Ind-Evaluate (n = 2)	0 (0)	0 (0)	100 (2)	50 (2)	0 (0)	0 (0)	50 (1)	50 (1)	100 (2)	0 (0)	50 (1)	50 (1)
Ind-General (n = 14)	0 (0)	0 (0)	92.9 (13)	64.3 (9)	42.9 (6)	42.9 (6)	50 (7)	42.9 (6)	50 (7)	14.3 (2)	100 (14)	78.6 (11)

Note: N in parentheses below percentage of participants. *VRIN-r* variable response inconsistency, *TRIN-r* true response inconsistency, *F-r* infrequent responses, *Fp-r* infrequent psychopathology responses, *Fs* infrequent somatic responses, *FBS-r* symptom validity, *RBS* response bias scale, *L-r* uncommon virtues, *K-r* adjustment validity.

Finally, after excluding non-content based invalid MMPI-2-RF results, Table 7 and Table 8 displayed how many participants scored above particular cutoff scores on each MMPI-2-RF validity scale after being reclassified on the SIRS-2. For those who remained in the Feigning category in both the civil and criminal samples, there is still clear evidence of infrequent responding and over-reporting of symptoms, which supports their classification of feigning. Results are similar for participants reclassified as Indeterminate-Evaluate or Indeterminate-General in both samples. Most importantly, 33.3% (n = 5) of participants reclassified as Honest in the civil

Table 6.

% Above Cutoff for MMPI-2-RF Validity Scale Results for 105 Criminal Defendants Classified as Feigning on the SIRS (Non-Content Based Invalid Results included)

	VRIN-r	TRIN-r	F-r	F-r	Fp-r	Fp-r	Fs	Fs	FBS-r	FBS-r	RBS	RBS
MMPI-2-RF Cutoff	80	80	100	120	90	100	90	100	90	100	90	100
SIRS-2 Classification												
Feigning (n = 52)	19.2 (10)	26.9 (14)	90.4 (47)	82.7 (43)	90.4 (47)	80.8 (42)	73.1 (38)	46.2 (24)	36.5 (19)	17.3 (9)	88.5 (46)	63.5 (33)
Genuine (n = 4)	0 (0)	25 (1)	100 (4)	75 (3)	75 (3)	50 (2)	0 (0)	0 (0)	0 (0)	0 (0)	50 (2)	25 (1)
Ind-Evaluate (n = 34)	17.6 (6)	11.8 (4)	91.2 (31)	85.3 (29)	85.3 (29)	70.6 (24)	76.5 (26)	38.2 (13)	32.4 (11)	20.6 (7)	88.2 (30)	76.5 (26)
Ind-General (n = 15)	20 (3)	33.3 (5)	80 (12)	73.3 (11)	73.3 (14)	73.3 (11)	53.3 (8)	33.3 (5)	6.7 (1)	0 (0)	66.7 (10)	60 (9)

*Note:* N in parentheses below percentage of participants. *VRIN-r* variable response inconsistency, *TRIN-r* true response inconsistency, *F-r* infrequent responses, *Fp-r* infrequent psychopathology responses, *Fs* infrequent somatic responses, *FBS-r* symptom validity, *RBS* response bias scale, *L-r* uncommon virtues, *K-r* adjustment validity.

sample yielded a score of 120 on the F-r scale, while 20.0% (n = 3) yielded a score of 100 on the Fp-r scale, which is representative of an invalid MMPI-2-RF protocol. Similarly, 66.7 % (n = 2) of participants in the criminal sample yielded a score of 120 on the F-r scale and 33.3% (n = 1) yielded a score of 100 on the Fp-r scale, despite their reclassification into the Honest group.

Table 7.

% Above Cutoff for MMPI-2-RF Validity Scale Results for 36 Disability Claimants Classified as Feigning on the SIRS (Non-Content Based Invalid Results removed)

	VRIN-r	TRIN-r	F-r	F-r	Fp-r	Fp-r	Fs	Fs	FBS-r	FBS-r	RBS	RBS
MMPI-2-RF Cutoff	80	80	100	120	90	100	90	100	90	100	90	100
SIRS-2 Classification												
Feigning (n = 1)	0 (0)	0 (0)	100 (1)	0 (0)	0 (0)	0 (0)	100 (1)	100 (1)	0 (0)	0 (0)	0 (0)	100 (1)
Genuine (n = 15)	0 (0)	0 (0)	86.7 (13)	33.3 (5)	20.0 (3)	13.3 (2)	40.0 (6)	26.7 (4)	40.0 (6)	26.7 (4)	86.7 (13)	66.7 (10)
Ind-Evaluate (n = 6)	0 (0)	0 (0)	100 (6)	83.3 (5)	0 (0)	0 (0)	66.7 (4)	33.3 (2)	100 (6)	16.7 (1)	83.3 (5)	66.7 (4)
Ind-General (n = 14)	0 (0)	0 (0)	92.9 (13)	64.3 (9)	42.9 (6)	42.9 (6)	50.0 (7)	42.9 (6)	50.0 (7)	14.3 (2)	100 (14)	78.6 (11)

*Note:* N in parentheses below percentage of participants. *VRIN-r* variable response inconsistency, *TRIN-r* true response inconsistency, *F-r* infrequent responses, *Fp-r* infrequent psychopathology responses, *Fs* infrequent somatic responses, *FBS-r* symptom validity, *RBS* response bias scale, *L-r* uncommon virtues, *K-r* adjustment validity.

Table 8.

% Above Cutoff for MMPI-2-RF Validity Scale Results for 69 Criminal Defendants Classified as Feigning on the SIRS (Non-Content Based Invalid Results removed)

	VRIN-r	TRIN-r	F-r	F-r	Fp-r	Fp-r	Fs	Fs	FBS-r	FBS-r	RBS	RBS
MMPI-2-RF Cutoff	80	80	100	120	90	100	90	100	90	100	90	100
SIRS-2 Classification												
Feigning (n = 32)	0 (0)	0 (0)	90.6 (29)	87.5 (28)	90.6 (29)	81.3 (26)	78.1 (25)	46.9 (15)	43.8 (14)	18.8 (6)	90.6 (29)	71.9 (23)
Genuine (n = 3)	0 (0)	0 (0)	100 (3)	66.7 (2)	66.7 (2)	33.3 (1)	0 (0)	0 (0)	0 (0)	0 (0)	33.3 (1)	0 (0)
Ind-Evaluate (n = 25)	0 (0)	0 (0)	88 (22)	84 (21)	88 (22)	72 (18)	80 (20)	40 (10)	32 (8)	20 (5)	84 (21)	80 (20)
Ind-General (n = 9)	0 (0)	0 (0)	66.7 (6)	55.6 (5)	55.6 (5)	55.6 (5)	33.3 (3)	11.1 (1)	0 (0)	0 (0)	55.6 (5)	55.6 (5)

*Note:* N in parentheses below percentage of participants. *VRIN-r* variable response inconsistency, *TRIN-r* true response inconsistency, *F-r* infrequent responses, *Fp-r* infrequent psychopathology responses, *Fs* infrequent somatic responses, *FBS-r* symptom validity, *RBS* response bias scale, *L-r* uncommon virtues, *K-r* adjustment validity.

## Chapter IV

### DISCUSSION

The authors of the SIRS-2 made significant revisions to the original SIRS with the goal of improving the widely used assessment. However, there have been numerous concerns about the development and validation of the SIRS-2, and even about its clinical utility in accurately classifying individuals. In the current study, SIRS and SIRS-2 feigning rates were examined in both civil and criminal forensic samples.

The results supported the first hypothesis that fewer participants would be classified as Feigning on the SIRS-2 than the SIRS in both samples. For the civil sample, about 14% (n = 5) of participants remained classified as Feigning on the SIRS-2 compared to 37 participants, who were in the same classification group on the original SIRS. For the criminal sample, about 50% (n = 52) of participants remained within the Feigning classification after 105 participants were classified in the same group on the original SIRS. It is clear that there was a significant decrease in the number of individuals who remained classified as Feigning on the SIRS-2 from the original SIRS in both samples. These results support the critical concern that the revisions to the SIRS-2 may have reduced the false positive rate, but at a significant cost to sensitivity since fewer individuals were identified as feigning. We also predicted that more participants would be reclassified in either Indeterminate category on the SIRS-2 if they did not fall into the Feigning classification. The remaining individuals in the civil forensic sample were reclassified as Honest (n = 16) or Indeterminate-General (n = 14), while only two participants were reclassified



as Indeterminate-Evaluate. The majority of the remaining individuals within the criminal sample were reclassified as Indeterminate-Evaluate (n =34) or Indeterminate-General (n = 15), while four participants were reclassified as Honest. Because the rate of feigning may increase in settings, such as correctional institutions, where individuals have less knowledge of severe psychopathology, it makes sense that more individuals within the criminal sample were reclassified as Indeterminate-Evaluate. The probability of feigning in this category exceeds 50%, as mentioned earlier. Additionally, criminal defendants yielded a higher mean score on the MMPI-2-RF Fp-r scale compared to disability claimants, which is consistent with previous research that found that the forensic context (i.e. criminal vs. civil) can influence the types of symptoms exaggerated during an evaluation (Wygant, et al., 2007). Since many items on the SIRS-2 inquire about symptoms of severe psychopathology, this result suggests that the SIRS-2 may be more applicable within correctional settings, where such symptoms are more prominent.

For individuals who did not remain in the Feigning classification on the SIRS-2, it was predicted that MMPI-2-RF validity scales would corroborate evidence of over-reporting and inconsistent responding. Once non-content based invalid results of the MMPI-2-RF were removed, the sample included 36 disability claimants. Of the five participants who were classified as Feigning before the invalid results were removed, only one participant remained classified as Feigning after the removal, which suggests that the SIRS-2 accurately identified individuals who are malingering psychopathology as Feigners. However, of the 15 participants who were classified as Honest on the SIRS-2, 86.7% (n = 13) of participants responded

infrequently on the MMPI-2-RF and five participants produced invalid protocols. After non-content based invalid results were removed in the criminal sample, 69 criminal defendants remained. Before the removal, 52 participants were classified as Feigning and 32 participants remained within that classification after the removal. Unlike the civil sample, the SIRS-2 was less accurate in identifying true feigners as feigners since the majority of criminal defendants within this classification produced invalid protocols even after the removal. Similarly, most criminal defendants in either Indeterminate category and the Honest categories also revealed evidence of over-reporting and infrequent responding after the removal of non-content based invalid results. Despite the classification that participants of both samples were grouped into using the SIRS-2, the current study examined collateral data that strongly suggested the participants were indeed over-reporting symptoms.

### **Limitations**

Although the current study highlights the weakness in clinical utility of the SIRS-2, the results should be considered in light of an important limitation. Participants within the criminal sample of the current study were only administered the SIRS-2 if evaluators suspected that they were feigning symptoms. This approach has led to a potential selection bias for the criminal sample, and therefore, the rate of feigning that has been observed may not be truly representative of the sample as a whole. However, this practice is also reflective of clinical assessment in the real world, such that evaluators only administer an assessment when examinees are suspected of experiencing particular symptoms, which can be assessed by a relative instrument. For that reason, the method used within this study may not actually be

such a limitation, even though the possibility of a sampling bias may be considered a weakness. Furthermore, the criminal sample only included male participants, which may have impacted the rate of feigning observed. Therefore, it is recommended that future research should include both males and females in their criminal samples.

### **Conclusion**

Since the release of the SIRS-2, there have been few studies to validate the measure. Various concerns about the clinical utility of the SIRS-2 remain given results found in past research and the current study. Since the SIRS was the most well-known and most frequently administered measure for feigned mental disorders, it was expected that the SIRS-2 would replace the original SIRS and remain as such. However, without further investigation and acknowledgement of the significant limitations of the SIRS-2, it may be too early to allow the SIRS-2 to supersede the SIRS.

The current study provides strong evidence that the sensitivity rate of the SIRS-2 has been reduced. Since identifying true feigners in forensic settings is extremely important for legal decisions and treatment recommendations, the lack in ability for the SIRS-2 to accurately identify such individuals should raise concern for examiners. It was also found that criminal defendants over-reported a higher number of symptoms of genuine, severe psychopathology than disability claimants. This suggests that the SIRS-2 may be more applicable in correctional settings, where individuals may feign symptoms of severe psychopathology. Future research should explore this finding in other correctional settings. Additionally, the SIRS-2 may be less useful in detecting feigners of particular symptoms in a civil setting.

Participants within the current civil sample were expected to feign more physical symptoms since they were involved in disability lawsuits. The SIRS-2 was less likely to classify these individuals as Feigners despite their high mean scores on the MMPI-2-RF symptom validity scales. This finding raises the question of whether the SIRS-2 should include additional items to accurately identify these types of feigners. If the SIRS-2 is assumed to replace the SIRS as the premier measure for feigned mental disorders, then the SIRS-2 should be able to detect feigners of various types of symptoms, including somatic, cognitive, and psychological.

## References

- American Psychiatric Association (2003). *Diagnostic and statistical manual of mental disorders* (5<sup>th</sup> ed.). Washington, DC: Author.
- Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment, 87*, 84-94.
- Ben-Porath, Y. S. (2003). Assessing personality and psychopathology with self-report inventories. In I. B. Weiner (Series Ed.) & A. M. Goldstein (Vol. Ed.), *Handbook of psychology: Vol. 11. Forensic psychology* (pp. 485–508). Hoboken, NJ: Wiley.
- Ben-Porath, Y. S., & Tellegen, A. (2008/2011). *Minnesota Multiphasic Personality Inventory-2-Restructured Form: Manual for administration, scoring and interpretation*. Minneapolis: University of Minnesota Press.
- 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.
- Burchett, D. L., & Bagby, R. M. (2014). Multimethod assessment of response distortion: Integrating data from interviews, collateral records, and standardized assessment tools. *Multimethod Clinical Assessment, 345-378*.
- DeClue, G. (2011). Harry Potter and the structured interview of reported symptoms. *Open Access Journal of Forensic Psychology, 3*, 1-18.
- Edens, J. F., Poythress, N. G., & Watkins-Clay, M. M. (2007). Detection of malingering in psychiatric unit and general population prison inmates: A comparison of the PAI, SIMS, and SIRS. *Journal of Personality Assessment, 88*, 33-42.
- Glassmire, D. M., Tarescavage, A. M., & Gottfried, E. D. (2016). Likelihood of obtaining structured interview of reported symptoms (SIRS) and SIRS-2 elevations among forensic psychiatric inpatients with screening elevations on the miller forensic assessment of symptoms test. *Psychological Assessment*. Advance online publication. <http://dx.doi.org/10.1037/pas0000289>
- Green, D., & Rosenfeld, B. (2011). Evaluating the gold standard: a review and meta-analysis of the Structured Interview of Reported Symptoms. *Psychological Assessment, 23*, 95.

- Green, D., Rosenfeld, B., & Belfi, B. (2012). New and improved? A comparison of the original and revised versions of the Structured Interview of Reported Symptoms. *Assessment, 20*, 210-218.
- Hathaway, S. R., & McKinley, J. C. (1943). *The Minnesota Multiphasic Personality Inventory*. Minneapolis: University of Minnesota Press.
- Kocsis, R. N. (2011). The structured interview of reported symptoms 2nd edition (SIRS-2): The new benchmark towards the assessment of malingering. *Journal of Forensic Psychology Practice, 11*, 73-81.
- Miller, H. A. (2001). *Miller-Forensic Assessment of Symptoms Test professional manual*. Odessa, FL: Psychological Assessment Resources.
- Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology, 24*, 1094-1102.
- Morey, L. (1991). *Professional manual for the Personality Assessment Inventory (PAI)*. Odessa, FL: Psychological Assessment Resources.
- Morey, L. (1996). *An interpretive guide to the Personality Assessment Inventory (PAI)*. Odessa, FL: Psychological Assessment Resources.
- Rogers, R., Bagby, R. M., & Dickens, S. E. (1992). *Structured Interview of Reported Symptoms (SIRS) and professional manual*. Odessa: Psychological Assessment Resources, Inc.
- Rogers, R., Gillis, J. R., Dickens, S. E., & Bagby, R. M. (1991). Standardized assessment of malingering: Validation of the Structured Interview of Reported Symptoms. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3*, 89.
- Rogers, R., Payne, J. W., Berry, D. T., & Granacher Jr, R. P. (2009). Use of the SIRS in compensation cases: An examination of its validity and generalizability. *Law and Human Behavior, 33*, 213.
- Rogers, R., Salekin, R. T., Sewell, K. W., Goldstein, A., & Leonard, K. (1998). A comparison of forensic and nonforensic malingerers: A prototypical analysis of explanatory models. *Law and human behavior, 22*, 353.
- Rogers, R., Sewell, K. W., & Gillard, N. D. (2010). *Structured Interview of Reported Symptoms (SIRS), 2nd Edition, professional manual*. Lutz, FL: Psychological Assessment Resources, Inc.

- Rogers, R., Sewell, K. W., Morey, L. C., & Ustard, K. L. (1996). Detection of feigned mental disorders on the Personality Assessment Inventory: A discriminant analysis. *Journal of personality assessment*, 67, 629-640.
- Rubenzler, S. (2010). Review of the Structured Interview of Reported Symptoms-2 (SIRS-2). *Open Access Journal of Forensic Psychology*, 2, 273-286.
- Smith, G. P., & Burger, G. K. (1997). Detection of malingering: validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law Online*, 25, 183-189.
- Tyner, E. A., & Frederick, R. I. (2013). Rates of computational errors for scoring the SIRS primary scales. *Psychological Assessment*, 25, 1367.
- Wygant, D. B., & Granacher, R. P. (2015). Assessment of validity and response bias in neuropsychiatric evaluations. *NeuroRehabilitation*, 36, 427-438.
- Wygant, D. B., Sellbom, M., Ben-Porath, Y. S., Stafford, K. P., Freeman, D. B., & Heilbronner, R. L. (2007). The relation between symptom validity testing and MMPI-2 scores as a function of forensic evaluation context. *Archives of Clinical Neuropsychology*, 22, 489-499.