ASSESSMENT OF MALINGERING IN A JAIL REFERRAL POPULATION:
SCREENING AND COMPREHENSIVE EVALUATION

DISSERTATION

Presented to the Graduate Council of the
University of North Texas in Partial
Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

By

Karen L. Ustad, B.S., M.S.
Denton, Texas
August, 1997

Psychological assessment of mentally disordered offenders requires a systematic consideration of response styles, including malingering and defensiveness. Important components of these evaluations are standardized diagnostic interviews. However, the ability of offenders to feign mental disorders on such measures to achieve such external incentives as treatment, placement on safer units, or possible release from jail remains uninvestigated. With a known-groups comparison with the data from the Structured Interview of Reported Symptoms as a criterion, 24 suspected malingerers were compared to 64 genuine patients on the Schedule of Affective Disorders and Schizophrenia (SADS), the abbreviated SADS-C, the Suicide Probability Scale, and the Referral Decision Scale. As expected, nearly all (96%) of the suspected malingerers were able to successfully feign on these measures. To improve the clinical assessment of malingering, decision rules for the SADS proposed by Rogers (1988, 1997) were tested for (a) screening persons for possible feigning and (b) rendering a clinical determination.
of feigning. These decision rules were generally confirmed. In addition, the usefulness of the SADS-C was evaluated for the first time. Initial results were promising with a positive predictive power of .84 and a negative predictive power of .95. Finally, a screening battery is proposed for further investigation that would allow for the rapid assessment of persons in need of immediate evaluation for severe mental disorders or possible malingering.
ASSESSMENT OF MALINGERING IN A JAIL REFERRAL POPULATION: SCREENING AND COMPREHENSIVE EVALUATION

DISSERTATION

Presented to the Graduate Council of the University of North Texas in Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

By

Karen L. Ustad, B.S., M.S.
Denton, Texas
August, 1997
TABLE OF CONTENTS

LIST OF TABLES ........................................... v
LIST OF FIGURES ......................................... vii

I. INTRODUCTION ........................................ 1
   Models of Malingering
      The Pathogenic Model
      The Criminological Model
      The Adaptational Model
   Assessment of Malingering
      Clinical Interview
      Specialized Measures of Malingering and Dissimulation
      The Schedule of Affective Disorders and Schizophrenia (SADS)
   Psychological Assessment in a Jail Setting
      Background
      Regulation of Jail Mental Health Services: The Need for Identification and Treatment of Mental Illness
      Programmatic Considerations
      Jail Suicide
   The Tarrant County Jail Project
      Delineation of Comparison Groups
      Development of Research Questions and Hypotheses

II. METHOD ................................................. 45
   Participants
   Materials
   Procedure

III. RESULTS ................................................ 60
   Reliability of the Measures
   Malingering on Diagnostic Screening Instruments
   Consistency Between the SADS and the SADS-C
   Convergent Validity Between the SIRS and the M Test, SADS and SADS-C
   Convergent Validity of the M Test and the SIRS
Convergent Validity between the SADS and the SIRS
Convergent Validity between the SADS-C and the SIRS
Identification of Malingering Scales on the SADS and the SADS-C
Scale Development for the SADS
Scale Development for the SADS-C
Development of a Screening Battery for Malingering and Mental Illness

IV. DISCUSSION ................................................. 93
Dissimulation on Diagnostic Screening Instruments
Screening for Suicidality
Assessment of Malingering
Inconsistent Responding
Concurrent Validity of SADS Models of Malingering
Malingering Scales on the SADS and SADS-C
The M Test: A Viable Screening Measure
Developing a Screening Battery for Malingering and Mental Illness
Limitations to the Current Study

NOTES ...................................................... 121
APPENDICES ................................................. 125
REFERENCES .............................................. 139
LIST OF TABLES

1. SADS Threshold and Clinical Decision Models of Malingering Proposed by Rogers (1988) ........................................ 27
2. Demographic and Background Characteristics of the Sample ......................................................... 43
3. Reliability of Psychodiagnostic and Screening Instruments ................................................................. 57
4. Means, Standard Deviations and Significance Levels for the RDS ......................................................... 59
5. Means, Standard Deviations, and Significance Levels for the SADS-C ..................................................... 60
6. Means, Standard Deviations, and Significance Levels for the SPS ......................................................... 61
7. Means and Standard Deviations of Difference Scores on the SADS and SAD-C by Group ....................... 64
8. Consistent Responding by Group on the Scales of the SADS-C ............................................................. 67
9. Correlations Between the SIRS and the M Test .................................................................................. 69
10. Predictive Accuracy of the M Test in the Detection of Malingering - Option A ........................................ 71
11. Predictive Accuracy of the M Test in the Detection of Malingering - Option B ....................................... 72
12. Convergent Validity Between the SADS-Part 1 and the SIRS .............................................................. 74
13. Discriminant Function Analysis for the SADS-C Decision Model ......................................................... 78
14. Means, Standard Deviations, and Significance Levels for the SADS ....................................................... 82
15. Predictive Accuracy of the SADS-Sum ......................................................................................... 83
16. Predictive Accuracy of the SADS-Count ......................................................................................... 84
17. ANOVA: Patient versus Suspected Malingerer Groups on the SADS-C-Sum and SADS-C-Count ........................................... 87

18. Predictive Accuracy of the SADS-C-Sum by Group - Threshold Model ................................................................. 89

19. Classification Accuracy of the SADS-C-Sum by Group - Clinical Decision Model ......................................................... 90

20. Classification Accuracy of SADS Models of Malingering .......................................................................................... 109
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Screening Battery for Malingering and Mental Illness</td>
<td>94</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

A critical issue for all mental health professionals is the accurate clinical assessment of a client's past and present psychological functioning. In mental health practice, the accuracy of clinical evaluations is complicated by the lack of purely objective measures of distress or dysfunction. Instead, clinicians rely to a great extent on self report measures, clinical interviews, and behavioral observation to assess the status of each client. The extent to which these assessments represent "reality" is, in part, restricted by the veracity and comprehensiveness of client's self report.

While issues of malingering and dissimulation are important in the context of any clinical evaluation, these response styles assume paramount importance when strong motivation to feign is apparent. According to Rogers (1987), the likelihood of guardedness or malingering increases in situations where patients view the circumstances to be adversarial and believe that the clinician's role may be at variance with their own needs. In forensic settings, the motivation to malinger may occur in evaluations where the stakes are high. Evaluations of
criminal responsibility, competency to stand trial, competency to be sentenced, and competency to be executed are examples of such instances (Adelman & Howard, 1984). These evaluations are adversarial in that mental health professionals are often consultants to the court and thus, may not satisfy or even address the individual needs of the patient (Rogers & Cavanaugh, 1983).

Malingering is defined as the intentional production of false symptoms and/or the exaggeration of true symptoms in pursuit of a goal that is recognizable in light of the individual's circumstances (American Psychiatric Association, 1994). While malingering can be conceptualized as a dichotomous variable (i.e., present or absent), it is perhaps more accurate to view this behavior along a continuum (Rogers, 1988). Individuals vary along a number of dimensions including degree of intentionality, degree of distortion, and motivation to malinger. Comprehensive assessment involves a thorough exploration of all components.

The following literature review begins with a presentation of the current conceptualizations regarding etiology and how this is related to motivation and incentive. The next major section provides an overview of current assessment instruments used in the detection of malingering with an emphasis on self-report and interview based methods. These measures are relevant to both forensic
and non-forensic evaluations. The third major section examines malingering within the specific context of correctional facilities in which a premium is placed on the rapid and accurate assessment of psychopathology. The emphasis in this final section is on the screening and comprehensive evaluation of jail detainees via structured interviews.

The terms possible and suspected malingerer will be used throughout and should not be considered synonymous. Possible malingering refers to situations in which the potential for malingering exists but there is no estimate on the likelihood that it did or did not occur. Suspected malingering, on the other hand, refers to situations in which there is a high likelihood that malingering has occurred and this is corroborated by an external criterion.

Models of Malingering

Motivation to malinger becomes an important consideration with respect to classification. For example, the differentiation between malingering and factitious disorder requires an understanding of the individual's motivation. More explicitly, the motivation for malingering is an external incentive whereas the motivation for factitious disorder is the adoption of a "sick role" (APA, 1994). Interestingly, individuals who feign for an external incentive are usually assumed to be socially deviant or "bad" and thus, should be punished for their dishonesty
(Rogers, 1990b). This assumption has not typically been applied to individuals feigning to assume a sick role. In an attempt to address the issue of motivation and etiology, Rogers (1990a; 1990b) outlined two existing models of malingering (the pathogenic and the criminological) and proposed a third (the adaptational). Rogers suggested that the adaptational model provides a conceptualization that is less judgmental and more ecologically valid than the two previous models.

The Pathogenic Model

The pathogenic model assumes that the patient is motivated by underlying psychopathology. Although initially under voluntary control, fabricated symptoms gradually devolve into a genuine disorder. Historically, this response style has been referred to as pseudomalingering. The notion of pseudomalingering dates back to the published works of Leonard Andreyev's (1902) novel, The Dilemma. In this novel, a physician intentionally committed a murder with the deliberate plan to feign insanity at the time of the offense. Later, the physician is surprised to find out that the feigned symptoms had become reality in the form of auditory hallucinations. Pseudomalingering is the feigning of a mental illness which later becomes real.

The pathogenic model proposes that malingering is an "ineffectual attempt to control psychotic and neurotic processes by consciously reproducing underlying pathology"
(Rogers, 1990b, p. 324). In other words, the model assumes that malingering is an early or prodromal stage of mental disease. This view was supported by many of the early dynamic theorists who proposed that malingering signified an arrest of development at an early phase (Eissler, 1951; Menninger, 1935). As discussed by Rogers (1990b), this model has lost its general acceptance with three major reasons for its decline: (a) current research has not supported the deteriorative trend from malingering to mental illness; (b) external, conscious motives for malingering are often identified; and (c) psychiatric facilities, once viewed as asylums to avoid, may offer tangible advantages, particularly over incarceration. In general, the pathogenic model has little utility in the study of malingering since it provides minimal information with respect to etiology or differential diagnosis.

The Criminological Model

The criminological model assumes that feigning mental illness is a purposeful behavior aimed at the acquisition of unwarranted rewards (i.e., better living quarters, financial gain) and does not represent the prodromal phase of a genuine disorder. This conceptualization is found in the definition of malingering as put forth in the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; American Psychiatric Association, 1994).
According to the DSM-IV, malingering is defined as a non-pathogenic response style that involves the intentional falsification or exaggeration of physical or psychological symptoms for the purpose of obtaining an external reward. The presence of an Axis I disorder does not preclude an individual from malingering current symptomatology and associated functional impairments. The American Psychiatric Association (1987; 1994) recommended that malingering should be suspected if any combination of the following is noted:

1. Medicolegal context of presentation (e.g., the person is referred by an attorney to the clinician for examination).
2. Marked discrepancy between the person's claimed stress or disability and the objective findings.
3. Lack of cooperation during the diagnostic evaluation and in complying with the prescribed treatment regimen.

The descriptive nature of this definition, while in accordance with the atheoretical stance of the DSM-IV, does not provide an explanation for why malingering occurs nor does it provide a good description of the discriminating characteristics of a malingerer. For example, the inclusion of APD as one of the potential "flags" of malingering, only
adds to the confusion of describing who might malinger. Rogers and Dion (1991) evaluated the number of possible permutations of APD and found that the combined DSM-III and DSM-III-R criteria yield a startling 29 trillion possible variations of the disorder. Furthermore, according to Rogers (1990a, 1990b), implicit in the DSM-III-R/DSM-IV model is the notion of "badness" and criminality. Thus, in Rogers categorization, the DSM-III-R/DSM-IV model is labelled the "criminologic" model.

The Adaptational Model

The adaptational approach, as applied to the forensic setting, considers malingering to be an active attempt to (a) circumvent involvement in the legal system, (b) provide a less punitive placement, or (c) lessen a defendant's culpability (Resnick, 1988). The adaptational model regards the process of malingering as the essential element in understanding this response style. The process is regarded as adaptive in that decisions are made on the basis of perceived outcome. Malingerers attempt to maximize their chances of achieving their desired goal(s) (Rogers 1990a, 1990b). The theoretical underpinnings of the model originate from research in decision making in which choices made under adversarial or uncertain conditions are based on utility and likelihood (Tversky & Kahneman, 1982; Von Neumann & Morgenstein, 1944).
A number of studies have been cited in support of the adaptational model (e.g., Braginsky, Braginsky, & Ring, 1969; Geller, Erlin, Kaye, & Fisher, 1990; Walters, White, & Greene, 1988). For example, Walters (1988) compared MMPI profiles of federal inmates under three conditions: neutral (no incentive to feign), adversarial/fake-bad (i.e., feign mental illness), and adversarial/fake-good (i.e., deny mental illness). The two adversarial conditions differed on the basis of incentive. The incentive in the adversarial/fake-bad condition was highly sought single-cell living; obtaining this incentive required that the inmate successfully present as mentally ill on the MMPI. Alternatively, the incentive in the adversarial/fake-good condition was a successful parole assessment; obtaining this goal required that the inmate produce a clinically healthy MMPI profile. The results of this descriptive study suggested that inmates under adversarial conditions can adjust their responses in the desired direction producing either a fake-bad or a fake-good profile. As expected, the MMPI validity indicators were found to be unremarkable in the neutral condition.

Rogers (1990a) also cited two other studies as preliminary support for the adaptational model. Rogers, Gillis, Dickens, and Bagby (1989) found that malingerers tend to be charged with serious offenses while non-malingerers tend to be charged with minor offenses. The
Implication is that a defendant faced with serious consequences is more likely to feign illness in an attempt to reduce the likelihood of experiencing negative or unwanted consequences. Similarly, the results of a quasi-naturalistic study conducted by Wilcox and Krasnoff (1967) indicated that the adversarial nature of the evaluation resulted in the production of fake-bad profiles on the Minnesota Multiphasic Personality Inventory (MMPI).

Rogers, Sewell, and Goldstein (1994) evaluated the validity of the three models using prototypical analysis. The authors compiled a list of the attributes of each model (8 pathogenic, 16 criminological, and 8 adaptational) and randomized their placement on a list. Forensic psychologists were asked to rate each item on its importance to the evaluation of malingering. The results suggested that attributes of the criminological and adaptational models are of greater importance than the pathogenic in the assessment of malingering. In addition, the authors assessed the validity of the models using principal component factor analysis with varimax rotation. A three factor solution produced distinct dimensions with minimal cross-loadings. These dimensions corresponded well to the theory-driven models proposed by Rogers and accounted for a total of 41.2% of the variance.
Assessment of Malingering

Clinical Interview

To date, much of the research related to malingering has focused on the ability of clinicians to detect malingering using clinical interviews and traditional psychological tests. Toward this end, Drob and Berger (1987) promoted a combination of methods in their proposed clinical-forensic model for the assessment of malingering. In their discussion of the model, the authors stressed the importance of both clinical interviews and psychological testing in the differential classification of malingering from bona fide mental disorders.

The clinical-forensic model of malingering is centered around three basic questions: (a) Does the patient exhibit the "classic signs" of feigned psychological symptoms? (b) Does the patient have a conscious motive or reward for appearing mentally ill? (c) Does the patient suffer from any actual mental defect or illness which would cause him to produce what would appear to be voluntary symptomatology? (p. 521).

A review by Resnick (1984) provided a detailed description of the "classic signs" of malingering that have evolved through the years (e.g., Davidson, 1949; Ossipov, 1944; Rogers, 1987; Ritson & Forest, 1970; Sierles, 1984; Wachpress, Berenberg, & Jacobson, 1953). Included among the 16 signs of malingered psychosis are items such as
malingers (a) may overact their part, (b) may be eager to call attention to their illnesses, (c) may report symptoms that fit no known diagnostic entity, and (d) are unlikely to show residual signs of schizophrenia (pp. 31-32). According to Resnick, the list is not exhaustive but can be useful in the identification of cases where evaluation of response style appears to be warranted.

Drob and Berger (1987) observed that clinical signs of malingering are best assessed through a combination of clinical interview, psychological testing, and ancillary sources. However, many of the "classic signs" cannot be evaluated using traditional psychometric testing since many are identified through interaction with and observation of the individual. Thus, the onus is on the clinician to accurately assess malingering. As stated in Drob and Berger's article, "we [clinicians] must at times act with some cunning, acting more like a detective than is typical in most psychiatric settings" (1987, p. 523). In contrast, Rogers (1988) provided a cogent argument against the over reliance on unstructured clinical and diagnostic interviewing in the assessment of malingering. Not only does the heterogeneity of interviews and interview styles preclude their standardization but the clinicians' ability to accurately identify malingers has repeatedly been questioned (Johnson, Klinger, & Williams, 1977; Rogers, 1984; Rosenhan, 1973, 1975).
The second question addressed by the clinical-forensic approach to malingering (Drob & Berger, 1987) is whether the clinician can detect the "conscious motive" underlying the feigned illness. The authors provide no direct guidelines for assessing motive but instead note that "establishing a motive is frequently a straightforward procedure which follows from an evaluation of the individual's circumstances and what he says about them" (p. 529). This highly inferential strategy appears to be biased against the client in that the authors also state that a person's statements are always suspect (i.e., denials of malingering should always be questioned) and that a clear admission of malingering may simply be a ploy to get the examiner to rule out this possibility.

Resnick (1984) also argued that evaluation of an individual's motivation to malinger is an important component of the assessment process. He stated that malingering is more likely to occur in situations where the incentive is high and often readily apparent. Specifically, Resnick provided a descriptive taxonomy used for the identification of circumstances where malingering is considered to be likely. These situations are as follows: (a) where there is incentive to avoid punishment, (b) where there is incentive to avoid undesired duties or obligations, (c) where financial gain is possible, and (d) where improvement of current life circumstances is possible. For
example, motivation to feign mental illness in competency evaluations likely falls into two categories: the avoidance of punishment (e.g., delaying trial proceedings indefinitely or having the charges dropped) or the improvement of current life circumstances (e.g., residing at a more comfortable inpatient forensic facility rather than a county jail). While not explaining the motivation for malingering, the descriptive taxonomy is useful in that it illustrates the importance of identifying the specific incentive associated with feigned mental illness.

The third question included in the clinical-forensic approach specifically addressed the differential diagnosis of malingering from true mental illness. The discussion provided by Drob and Berger (1987) is conspicuously lacking in that they fail to mention existing empirical methods used to assess psychopathology and malingering. Instead, the authors provided a decidedly circular argument which stated that a person who evidences many of the "classic signs" of malingering [question 1] . . . and/or incentive to feign illness [question 2] should be suspected of malingering [question 3]. While not discounting the importance of the differential diagnosis of malingering, the authors do not appear to have fully addressed the available strategies for discriminating between authentic and feigned mental illness. In fact, much of the research in the assessment of malingering has involved indices on multiscale personality
inventories (Greene, 1988). For example, malingering and/or dissimulation indices are incorporated on the California Psychological Inventory (Megargee, 1972), the Millon Clinical Multiaxial Inventories, I, II, and III (Millon, 1983, 1987, 1995), the Basic Personality Inventory (Jackson, 1989), the 16 Personality Factor Questionnaire (Karson & O'Dell, 1976), and the Minnesota Multiphasic Personality Inventories (MMPI and MMPI-2; Hathaway & McKinley, 1943; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989). The literature pertaining to malingering on multiscale inventories is extensive (especially the MMPI and MMPI-2); for recent reviews please see Green (1997), Rogers, Sewell and Salekin (1997) and Ustad and Rogers (1996). Because these reviews are peripheral to the current topic, they will not be addressed in this literature review.

Specialized Measures of Malingering and Dissimulation

Researchers have supplemented dissimulation indices on multiscale inventories with the development of measures designed specifically to assess malingering. At present, nine clinical instruments have been developed that are specific to the assessment of malingering: the Malingering Scale (MS; Schretlen, 1988; Schretlen & Arkowitz, 1990), the M Test (Beaber, Marston, Michelli, & Mills, 1985), the Structured Inventory of Malingered Symptoms (SIMS; Smith & Burger, 1993), the Structured Interview of Reported Symptoms (SIRS; Rogers, Gillis & Dickens, 1992), Tehachapi
Malingering Scale (TMS; Haskett, 1995), the Malingering Detection Scale (MDS; Barkemeyer & Callon, 1989), the Malingering Probability Scale (MPS; Silverton & Gruber, in press), the Inventory of Problems (IOP; Viglione & Landis, 1995), and the Sentence Completion Test (SCT-136; Timmons, Lanyon, Almer, & Curran, 1993). Of these, the SIRS is the only standardized instrument for the detection of feigned mental disorders that is commercially available to mental health professionals. The following two sections provide detailed descriptions of the M Test and the SIRS. At present, these are the two most commonly studied instruments used in the assessment of feigned psychopathology.

**M Test**

The M Test is a 33-item, true-false, paper and pencil test that is designed to identify the malingering of schizophrenic symptoms (Beaber et al., 1985). In its original form, the M Test was composed of three scales: the Confusion (C), Schizophrenia (S), and Malingering (M) scales. Item selection was nonempirical in that items were chosen on the basis of their relationship to other rationally derived scales. Specifically, the C scale is composed of eight items that reflect beliefs and attitudes that are not symptoms of mental illness. The S scale is composed of ten items that are true indicators of schizophrenia as identified in the DSM-III (American Psychiatric Association, 1980). The M scale is composed of
15 items that represent fictitious symptoms in the form of (a) atypical hallucinations and delusions, and (b) symptoms of extreme severity.

Initial validation of the M Test was hampered by methodological shortcomings such as a complete reliance on a simulation design and the use of an all male sample (Rogers, Bagby, & Gillis, 1992). In spite of these limitations, the results of the initial validation study were considered to be promising. For example, a series of chi-square analyses revealed that the M Test accurately identified 78.2% of male students feigning schizophrenia and 87.2% of the schizophrenic inpatients (Beaber et al., 1985). The cut scores utilized in this study were a score of three or more on the C scale and/or a score of four or more on the M scale. Individuals who met either of these criteria would be recommended for further evaluation for possible malingering. With respect to normative data, Beaber et al. reported that items on the M Test were endorsed less than 10% of the time by a non-patient, college sample.

Recent studies using the M Test have shown conflicting results regarding its effectiveness as a screening measure for malingering (Gillis, Rogers, & Bagby, 1991; Hankins, Barnard, & Robbins, 1993; Schretlen, Neal, & Hochman, 1995; Smith & Borum, 1992). Gillis and colleagues (1991) evaluated the predictive validity of the M Test in a study which utilized both known groups and simulation design
methodology. Similar to the results of Beaber et al. (1985), the M Test was able to accurately classify 79.8% of the simulators feigning schizophrenia. However, the classificatory accuracy of the M Test dropped significantly when evaluating individuals suspected of malingering. Specifically, only 40% of the suspected malingerers were identified using Beaber's cut score for the M scale. Gillis and colleagues also evaluated the accuracy of classification between the two most clinically important groups: suspected malingerers and honest responders diagnosed with an Axis I disorder. The results of this analysis were very discouraging since only 28% of the suspected malingerers were accurately classified. In light of this low sensitivity, the authors suggested that the M Test, at least in its original form, is not well-suited as a screen for malingering.

Rogers, Bagby, and Gillis (1992) re-evaluated the data obtained in the Gillis et al. (1991) study with the objective of improving the effectiveness of the M Test. Based on the computation of positive predictive power and negative predictive power, the authors were able to create two distinct scales each having ten items. These scales were labelled the Rule-Out scale and the Rule-In scale. The Rule-Out scale was designed to eliminate bona fide patients while the Rule-In scale was designed to identify potential malingerers. The internal consistency of each scale was
reported to be excellent (Rule-Out scale, alpha = .85; Rule-In scale, alpha = .87).

Given the different priorities of clinical versus forensic settings, Rogers et al. (1992) recommended that different cut scores be used in each setting. Since malingering is likely to be a greater concern in forensic settings, a cut score of zero for the Rule-In scale was recommended (i.e., scores of ≥ 1 suggest possible malingering). The classification rate obtained with this cut score was excellent with 95.5% of the suspected malingerers being accurately classified. Cut scores for the clinical setting were less stringent on the Rule-In scale (i.e., less than two) and this was reflected in the slightly lower classification rate of 83.8%. The recommended cut score for the Rule-Out scale remained the same regardless of the population ( < 4). The authors concluded that the Rule-In and Rule-Out scales provide a quick and effective way of screening out bona fide patients where malingering of Axis I disorders may be an issue.

A recent attempt to cross-validate the Rule-In and Rule-Out scales was marginally successful in that the scales were cross-validated but the classificatory accuracy was substantially lower (Smith, Borum, & Schinka, 1993). Using the most conservative rule (i.e., a score of zero on the Rule-In scale), accurate classification of malingerers
dropped to 72.7% while the false positive rate increased from 23.5% to 50.8%.

Hankins, Barnard, and Robbins (1993) questioned the validity of the new scales on the basis of methodological concerns they believed were present in the original validation study. The researchers expressed concern on the following four issues: (a) the malingering sample may have been too heterogeneous with respect to symptomatology to provide an adequate test of validity, (b) the study's high base rate of suspected malingering likely maximized the predictive power of the instrument, (c) the use of extreme cases of probable malingering may have exaggerated group differences and inflated results, and (d) cross-validation was necessary before the test could be recommended for use outside the realm of research. Ultimately, Hankins and colleagues provided at least minimal support for the utility of the Rule-In and Rule-Out scales. They noted that the measure may useful as a screen for malingered psychosis but they stress the need for further research before the measure can be recommended for routine clinical use. Given the equivocal findings, more research is necessary before the validity and the generalizability of the Rule-In and Rule-Out scales can be established.

Structured Interview of Reported Symptoms (SIRS)

The development of structured and semi-structured interviews has provided researchers and clinicians with a
systematic and reliable method of assessing psychological functioning (Rogers, Bagby et al., 1992). The advantages of this format include the standardization of content and style of clinical inquiries and the provision of a systematic procedure for quantifying clinical data (Rogers, 1987). It was with these goals in mind that the Structured Interview of Reported Symptoms was developed.

The SIRS is a structured interview designed to detect malingering and other related response styles. Interpretation involves eight primary scales: Rare Symptoms (RS), Symptom Combinations (SC), Improbable or Absurd Symptoms (IA), Blatant Symptoms (BL), Subtle Symptoms (SU), Severity of Symptoms (SEV), Selectivity of Symptoms (SEL), and Reported vs. Observed Symptoms (RO). Five supplementary scales are also included: Direct Appraisal of Honesty (DA), Defensive Symptoms (DS), Symptom Onset (SO), Overly Specified Symptoms (OS), and Inconsistency of Symptoms (INC).

Classifications are made using single scale or multiple scale scores that place the individual in one of three categories: Honest responding, indeterminate, and feigning. The test was constructed with the intent to minimize false positives such that the acceptable false positive rate was 0% for the validation studies. Thus, the likelihood of being misclassified as malingering by the SIRS is minimal. The authors state, "In comparison with the standard paper
and pencil measures the SIRS represents a substantial improvement in that (a) it has been validated using simulation and known groups design, (b) the optimal classification rates have been replicated many times, (c) it appears effective in the detection of coached simulators and (d) it has been shown effective in detecting feigned specific disorders" (Rogers, Bagby, & Dickens, 1992, p. 5).

To date, data have shown that the psychometric properties of the SIRS are satisfactory with respect to interrater reliability and internal consistency. Interrater reliability has been shown to be uniformly high across three recent studies (Linblad, 1993; Rogers, Gillis, Dickens, & Bagby, 1991; Rogers, Kropp, & Bagby, 1992). Interrater reliabilities range from .87 to 1.00 with mean reliabilities across studies of .95, .96 and .98. Each study differed with respect to the raters level of expertise; it appears that interrater reliability can be achieved by professionals and nonprofessionals who have received training and supervision in the administration of the SIRS (Rogers, 1996). Internal reliability was evaluated in the initial validation studies and found to be satisfactory with mean alpha coefficients of .86 for the primary scales and .75 for the supplementary scales (Rogers, Bagby, & Dickens, 1992).

Validation of the SIRS has involved multiple studies that have evaluated the construct using different
methodological approaches. In general, the validity of the SIRS has been supported by research utilizing simulation and known-groups designs (Rogers, 1997). Validation using simulation design has shown that the measure is effective in discriminating simulators from comparison groups, both clinical and non-clinical, across a range of settings with varied instructional sets (see, Connell, 1991; Rogers, et al., 1990; Rogers, Gillis, Bagby, & Monteiro, 1991; Rogers, Gillis, Dickens et al., 1991; Rogers, Kropp, & Bagby, 1993). For example, Kropp (1992) evaluated whether high levels of psychopathy improved an individual's ability to successfully feign mental illness. The results indicated levels of psychopathy, as measured by the Psychopathy Checklist-Screening Version (PCL-SV; Hare, Cox, & Hart, 1989), were unrelated to actual and perceived success.

Linblad (1993) conducted a within-subject, simulation design to evaluate whether experience as a psychiatric patient improved an individual's ability to feign mental illness. Comparison groups included non-mentally ill offenders, offenders with major mental illness (e.g., Axis I diagnoses), and offenders with characterological disorders (e.g., Axis II diagnoses). The SIRS accurately identified simulated versus bona fide profiles and was not affected by the presence of mental illness.

A study evaluating the effects of coaching on successful simulation found that coached simulators were
able to more successful at lowering their SIRS profiles when compared to uncoached counterparts (Rogers, Gillis, Bagby et al., 1991). It is important to note that the SIRS was able to accurately identify coached and uncoached simulators from bona fide patients. This finding suggests that coaching, in and of itself, does not necessarily help in the production of feigned mental illness.

The results of Kurtz and Meyers (1994) study support the findings of Rogers and colleagues. Kurtz and Meyers found that the SIRS was effective in discriminating between honest responders (i.e., inmates with bona fide mental illness) and coached simulators (i.e., inmates without a mental health history who were asked to feign psychosis). The classification rate of the SIRS (88.9%) and was superior to that of the MMPI-2 and the M Test at 81.3% and < 75%, respectively.

Rogers, Gillis, Dickens, et al. (1991) provide evidence of convergent and discriminant validity of the SIRS. In this study, the 12 SIRS scales were correlated with nine indicators of dissimulation found on the MMPI: L, F, K, F-K index, Gough's Dissimulation Scale - revised (Dsr; Gough, 1957), Weiner-Harmon Obvious Versus Subtle scales (Greene, 1980; Weiner, 1948), Lachar and Wrobel Critical items scale (Green, 1980; Lachar & Wrobel, 1979), Test-retest index (Buechley & Ball, 1952; Dahlstrom, Welsh, & Dahlstrom, 1972; Green, 1980) and the Carelessness scale (Green, 1978). In
general, the results showed that the SIRS scales were (a) positively correlated with MMPI fake-bad indices (e.g., 78.2% of the scales correlating at $\geq .60$), (b) negatively or nonsignificantly correlated with MMPI indices of defensiveness, and (c) modestly correlated with MMPI indices of inconsistency.

Convergent validity of the M Test and the SIRS was evaluated in the Rogers, Gillis, Dickens et al. (1991) study. Mean correlations between the M Test and the SIRS (excluding the DS scale) were .72 for the M scale, .67 for the C scale, and .65 for the S scale. These correlations are considerably higher than those of the original validation studies where mean correlations between the SIRS primary scales and the M Test were .51 for the C scale, .27 for the M scale, and .20 for the S scale (Rogers, Bagby, & Dickens, 1992). The results of both studies provide some support for the authors' initial hypotheses that (a) moderate correlations would be found between the SIRS and the M and the C scales (both scales of dissimulation), and (b) low correlations would be found between the SIRS and the S scale (a scale of bona fide schizophrenic symptomatology).

As already mentioned, the SIRS is the only structured interview designed specifically to assess malingering. However, Rogers (1987, 1988, 1997) suggested that the use of a general structured or semi-structured interview designed
to assess psychopathology could be helpful in the identification of suspected malingering.

The Schedule of Affective Disorders and Schizophrenia (SADS)

The SADS is a semi-structured interview developed for the differential diagnosis of mood disorders and schizophrenia (Spitzer & Endicott, 1978a). The interview is composed of two separate sections: Part 1 evaluates the client's most recent episode while Part 2 examines the client's psychiatric history. The interview is organized by three levels of inquiry: standard questions (i.e., inquiries asked of each patient), standard probes (i.e., established questions used when information from standard questions is not sufficient to make a rating), and optional probes (i.e., questions generated by the examiner which are used to extend the information base when responses to optional probes are not sufficient to make a rating). In comparison with traditional clinical interviews, the SADS provides the structure required to control for variability related to examiner effects (e.g., style or emphasis on questions) and therefore isolates variability to the client's self-report (Rogers, 1988).

One advantage of the SADS is its focus on identifying the severity and the duration of psychiatric symptomatology. Each item is rated by the examiner on a continuum, typically ranging from zero to six (i.e., absence of symptoms to
presence of severe symptoms). These ratings provide an advantage in the evaluation of malingering since clients are likely to find it more difficult to maintain a consistent symptom picture when the task requires recollection of severity ratings. Thus, maintaining a believable and consistent fabrication could be problematic for a malingerer when evaluated with the SADS.

**SADS Decision Models of Malingering**

Rogers (1988) reported SADS data on 104 patients that were court-referred for forensic evaluation. The data included 40 new cases and 64 cases that had previously been reported (see, Rogers, Thatcher, & Cavanaugh, 1984). From this sample, Rogers identified five clinical indicators of malingering: rare symptom endorsement, contradictory symptom pairs, unusual symptom combinations, symptom severity, and over-endorsement of symptoms. These indicators are described individually below:

1. Rare symptoms are those that occur infrequently in a psychiatric population (e.g., poverty of content of speech, neologisms). The scale includes ten items that were endorsed by less than 5% of the forensic sample with no bona fide patient reporting more than three (Rogers, 1988; Rogers, Thatcher, & Cavanaugh, 1984). All of the items on this scale relate to rare psychotic features suggesting that it may be most useful in identifying individuals who are trying to feign psychosis or global psychiatric impairment.
2. Contradictory symptom pairs include seven pairs of symptoms that measure psychological impairment in opposite directions (e.g., insomnia and hypersomnia, worthlessness and grandiosity). Endorsement of both items suggests the use of an inconsistent response style. Clinicians are encouraged to rule out rapid cycling bipolar disorder or cyclothymia prior to interpreting response style. The items on this scale are primarily related to affective disorders and may appear contradictory when the presence of such a disorder is overlooked. Rogers (1988) suggested that these symptom pairs are most useful in detecting unsophisticated malingerers who arbitrarily over-endorse psychiatric symptoms.

3. Unusual symptom combinations include pairs of common symptoms that are not contradictory to one another but are seldom found together in the same psychiatric patient. There are 12 symptom pairs that include one of the following three symptoms: appetite problems, psychomotor agitation, or persecutory delusions. These symptoms are then combined with one of the eight symptoms: current worrying, feelings of inadequacy, discouragement, psychic anxiety, anger, distrustfulness, or insomnia. The symptom combinations are based on actuarial data rather than a conceptual framework. Rogers (1988) suggested that these items are likely to have greater applicability with
sophisticated patients who may be able to recognize more obvious detection strategies.

4. Evaluation of symptom severity rests on the logic that feigners endorse too many symptoms as "extreme" in terms of frequency and resulting impairment. Using the scales of the SADS, Rogers developed cut scores for identifying individuals that employ this response style. Cut scores were developed on the basis of frequency: for symptoms rated $\geq 5$ the cut scores were 16 and 19 and for ratings $> 5$ the cut scores were 9 and 11.

5. Over-endorsement of symptoms refers to the extensive endorsement of symptoms in the clinical range (i.e., ratings of $\geq 3$). This response style is characterized by endorsement of numerous symptoms in a non-selective manner. In the 1988 study conducted by Rogers, forensic patients reported an average of 24.2 ($SD = 11.3$) symptoms in the clinical range with less than 5% of the sample reporting 45 or more symptoms and less than 1% reporting 48 or more symptoms.

On the basis of the normative data available, Rogers created two models of malingering: The Threshold Model and the Clinical Decision Model. The Threshold Model has relatively lenient inclusion criteria and was developed to maximize sensitivity and negative predictive power. In contrast, the criteria of the Clinical Decision Model is quite stringent and designed to maximize specificity and
positive predictive power (see Table 1 for criteria of both models and Appendix A for a listing of items by scale).

Table 1
SADS Threshold and Clinical Decision Models of Malingering Proposed by Rogers (1988)

I. Threshold Model of malingering in SADS evaluations

Any of the following criteria:
1. Rare symptoms ≥ 2
2. Contradictory symptoms ≥ 1
3. Unusual symptom combinations ≥ 1
4. Symptom severity ≥ 8 symptoms scored ≥ 5
5. Over-endorsement of symptoms ≥ 30 in the clinical range (≥ 3)
6. Marked inconsistencies between unstructured self-report and the SADS

II. Clinical Decision Model of malingering in SADS evaluations

A. Any of the following criteria:
1. Rare symptoms ≥ 3
2. Contradictory symptoms ≥ 3
3. Unusual Symptom combinations ≥ 5

(Table continues)
4. Symptom severity $\geq 17$ symptoms with scores 5 or more

5. Over-endorsement of symptoms $\geq 45$ symptoms in the clinical range ($\geq 3$)

B. Malingered presentation is corroborated by either:

1. Patient's admission of malingering

2. Clinical evidence of exaggeration and fabrication based on observation and/or psychometric data

To date, research on the SADS and malingering has not included the use of an independent criterion for malingering. The tentative model proposed by Rogers (1988) was developed normatively and involved the identification of very atypical response patterns. Validation of the clinical decision and Threshold Models of malingering can be initiated using traditional measures of malingering or the SIRS as the gold standard.

Psychological Assessment in a Jail Setting

Background

Over the past two decades, changes in mental health care and criminal justice procedures have radically altered the population characteristics of correctional facilities in the United States (Goldfarb, 1976; Lamb, Schock, Chen, & Gross, 1984; Ringel & Segal, 1986; Steadman, McCarty, & Morrissey, 1989; Teplin, 1984). For example, between 1978
and 1986 jail facilities encountered a 73% increase in population with many of these individuals being classified as mentally disordered (Steadman, McCarty, & Morrissey, 1989). The expansion of the psychiatric population in jail settings has been attributed to the process of deinstitutionalization resulting in a large number of mentally ill being diverted to community settings where they were unable to cope effectively with their behavior problems (Bachrach, 1980; Bassuk & Gerson, 1978; Steadman et al., 1989; Teplin, 1983). The question as to whether the process of deinstitutionalization was responsible for the criminalization of the mentally ill or simply led to a heightened awareness of an already existing problem is an important topic of debate (Steadman & Rubner, 1980). Although this question will not addressed in the current study, it is important to note that increased awareness led to the formation of commissions and task forces with the goal of developing mental health programs in correctional institutions (see, American Association of Correctional Psychologists, 1980; American Correctional Association, 1981; American Medical Association, 1981; American Psychiatric Association, 1974; 1989; American Public Health Association, 1976; Brodsky, 1983; Steadman et al., 1986).
Regulation of Jail Mental Health Services: The Need for Identification and Treatment of Mental Illness

Programmatic Considerations

Steadman, McCarty, and Morrissey (1986, 1989) conducted a national study of the mental health services provided at 43 jails in 26 states. The authors identified eight services that constituted core elements of mental health programming in local jails: (a) intake screening, (b) psychological evaluation, (c) competency evaluation, (d) medication evaluation, (e) substance abuse counseling, (f) psychotherapy, (g) inpatient care, and (h) hospitalization. Following from this survey, the authors provided five planning principles for the development of jail mental health programs. Relevant to the current study, the fourth principle stated that one of the facility's primary responsibilities to mentally ill jail detainees is the screening for mental health services. This screening includes suicide risk assessment and the identification of dangerous inmates (Roesch, 1991). While not specifically mentioned, the assessment of malingering is a crucial component in the mental health screening procedure. At the time of the Steadman project, only 70% of the jails had mental health screening as part of their program.
Jail Suicide

A five-year study in New York City correctional institutions found that 41% of all inmate deaths were suicides rendering suicide the leading cause of death in such institutions (Novick & Remlinger, 1978). Hayes and Rowan (1988) conducted a nation-wide study of jail suicides and found that 89% of the victims were pretrial detainees. The most critical time was the first 24 hours (Rakis, 1984; Rood & Faison, 1988) with most of the deaths occurring within the first three hours of incarceration (Hayes & Rowan, 1988). These studies illustrate the importance of prompt reliable assessment of suicide risk in jail settings.

Despite court mandated assessment and treatment of mentally disordered detainees (Bowring v. Godwin, 1977; Jones v. Wittenberg, 1980), as of 1988, most facilities did not provide mental health evaluation on intake (Teplin & Schwartz, 1989) and few reported having suicide prevention programs in place (Hayes & Rowan, 1988). Teplin and Schwartz stated that the absence of standardized procedures for identifying mentally disordered detainees may result in two problematic conditions. First, jail personnel may be unable to differentiate bona fide patients from individuals exhibiting disorderly conduct. Second, the social structure of the jail is such that an obedient, yet
mentally ill detainee, may be completely overlooked because their problems do not manifest in behavioral symptoms. For example, a detainee who is currently in a major depressive episode that presents in a quiet, withdrawn manner, may pose a high risk for suicide if treatment is not initiated in a timely manner. Despite the concern for accurate evaluation of suicidality; to date, empirical evaluation of feigned suicidality has not been conducted.

The advantages of screening for malingering at intake are in relation to appropriate distribution of mental health services and the costs associated with misappropriation of services. The key to effective screening is to provide a brief evaluation that identifies detainees who are in need of an extensive diagnostic workup and detainees in need of special precautions to protect themselves or others from harm (Dvoskin, 1989; Teplin & Schwartz, 1989).

Screening instruments tend to be brief, easy to administer and are particularly useful in situations where time or mental health resources are scarce and evaluation is necessary. However, the benefits of a screening procedure are limited by its effectiveness in relation to positive predictive power (PPP) and negative predictive power (NPP). In the case of some screening instruments (e.g., the Referral Decision Scale [RDS]), brevity may
come at the expense of validity as it is likely to be susceptible to false positive and false negative classifications. Thus, an effective screening instrument would be one that maximizes PPP without a substantial reduction in NPP. At present, research is sparse on the diagnostic screening of malingering.

The Tarrant County Jail Project

The Tarrant County Jail Project was designed to address a variety of issues relevant to malingering in both forensic and non-forensic settings. The study was developed to address five main questions:

1. How does dissimulation affect diagnostic screening instruments?
2. Do malingerers respond inconsistently across test sessions?
3. Do existing measures of malingering show convergent validity with the SIRS?
4. Can effective malingering scales be developed for the SADS and the SADS-C?
5. Is it feasible to develop a screening battery for jail settings with the measures used in the current study?

Delineation of Comparison Groups

The comparison groups used throughout the study were delineated as patient, suspected malingerer and indeterminant. The entire sample was comprised of jail
detainees referred to the Department of Mental Health and Mental Retardation for evaluation of psychiatric disturbance and need for treatment.

**Patient Group**

The patient group was comprised of participants who did not show evidence of malingering on the SIRS (i.e., 6 to 8 scales in the honest range with remaining primary scale scores in the indeterminate range). According to initial validation studies, application of these criteria resulted in a 95% probability that no feigning had occurred.

**Suspected Malingerer Group**

The suspected malingerer group was comprised of participants who showed evidence of malingering on the SIRS. A participant was classified as a suspected malingerer in one of two ways: (a) three or more of the primary scales were in the probable range, or (b) any one of the primary scales in the definite range. According to Rogers, Bagby, and Dickens (1992), this decision rule resulted in a 98% probability that those who met or exceeded the cut score were feigning symptomatology.

**Indeterminant Group**

The indeterminant group was comprised of participants who did not meet the standards for either honest (patient group) or malingered responding. The very stringent inclusion rules for the patient and suspected malingerer
groups resulted in the formation of a heterogeneous indeterminant group. Based on original validation studies, the probability that participants in this group were actually honest responders ranged from 18.2% to 72.2% while the probability that they were feigners ranged from 27.8% to 81.8%. For the purposes of the current project, statistical comparisons were conducted using either of the following groupings: (a) suspected malingerer and patient groups and (b) suspected malingerer, patient, and indeterminant groups.

Development of Research Questions and Hypotheses

Dissimulation on Diagnostic Screening Instruments

The effects of dissimulation were evaluated on a select group of diagnostic screening instruments: the RDS, SADS-C, and Suicide Probability Scale (SPS). Investigation of dissimulation on screening instruments is important because screening is often the first step in the identification of inmates in need of mental health services (see, Lamb, Schock, Chen, & Gross, 1984; Ogloff, Tien, Roesch, & Eaves, 1991; Pogrebin, 1985; Roesch, Ogloff, & Eaves, 1995). While the use of standardized measures is not universal in jail settings, these instruments have the benefit of being time effective, easy to administer and score, and have been shown to have modest to excellent discriminative ability. However, the brevity and the high face validity of most screening
instruments render them susceptible to feigned responding. Moreover, unlike more comprehensive instruments (e.g., the MMPI-2, PAI), these scales do not have direct means for evaluating dissimulation.

Thus, evaluation of dissimulation on screening instruments provides information on how these instruments are affected by a feigned response style. As previously described, malingered response styles are often characterized by an overendorsement of symptoms and severe ratings of distress. It was hypothesized that the suspected malingerer group would produce significantly higher scores than patients on all three measures and that they would screen positive for suspected mental illness.

Inconsistent Responding Across Time

As already mentioned, one of the classic signs of malingering is inconsistent reporting of symptoms (e.g., Rogers, 1984). The current study evaluated the consistency of self-report by comparing severity of symptom endorsement on the SADS-C to parallel items on the SADS-Part 1. The assumption underlying this comparison is that it is more difficult for malingerers, as compared to genuine patients, to maintain a convincing and consistent presentation over time. If this assumption is true, then the pattern of symptom endorsement between the SADS-C and the SADS would be more consistent for patients than suspected malingerers.
The SADS-C and the SADS were considered to be particularly good measures for addressing the question of response consistency. The standardized administration of these measures allows the clinician to attribute inconsistent responding to differences in patient's self-report rather than changes in style/emphasis of questions or idiosyncratic reporting of information (Rogers, 1988).

**Convergent Validity between the SIRS and the M Test, SADS and SADS-C**

**M Test and SIRS.** To date, research has been equivocal with respect to the validity of the M Test as a screen for malingering. The SIRS, on the other hand, has received extensive validation and is commercially available as a measure of malingering and other response styles. Evaluation of the convergent validity of the M Test and the SIRS is important to the current study as it provides an indication of its applicability for use in screening batteries. For the purposes of this study, the M Test was scored according to the two-step Rule-In/Rule-Out procedure (Rogers, Bagby, & Gillis, 1992).

**SADS and SIRS.** The concurrent validity of the SADS models of malingering were evaluated in relation to SIRS using suspected malingerer and patient groups for comparison. Both the SIRS and the SADS models were developed on the basis of prior research suggesting that malingerers can be detected on the basis of specific
strategies. Since the SADS models employ only a subset of the strategies included on the SIRS, it was anticipated that the convergence would be strong but that classification would not be identical.

In addition, the Clinical Decision Model and the Threshold Model vary on the type of misclassifications. The Clinical Decision Model was developed to be very conservative such that classification errors were in the direction of missing possible malingerers rather than misclassifying bona fide patients. The Threshold Model, on the other hand, is a more liberal model designed to identify the majority of possible malingerers. For the Threshold Model, errors in classification tend to involve the misclassification of bona fide patients as suspected malingerers. It was expected that classification would be in accordance with the design of the models; the Threshold Model would evidence high sensitivity and NPP while the Clinical Decision Model would evidence high specificity and PPP.

In addition, convergent validity of parallel scales between the SIRS (i.e., RS, SEL, and SEV) and the SADS (i.e., Rare Symptoms, Overendorsement of Symptoms, and Severe Symptom Endorsement; see Rogers, 1988) was investigated. Convergence between similar scales was expected since the scales were developed on the basis of the same strategy. However, the symptom severity scales
were expected to evidence only a modest correlation since the scoring procedures differ in an important way. Specifically, the severity measure on the SADS is composed of items rated by the examiner as 5 or higher whereas that of the SIRS is scored only when the examinee states that the symptoms are classified as "unbearable."

**SADS-C and SIRS.** The SADS Clinical Decision Model and the Threshold Model of malingering model were applied to the SADS-C. Based on the psychometric similarities between the SADS and the SADS-C and the theoretical basis of model development, it was assumed that the models would translate to the SADS-C without significant changes in efficacy. A finding of convergence would provide support for the construct validity of the models and would be an important step in developing a brief screen for malingered presentations that is embedded in a measure of current symptomatology.

The translation of the SADS models to the SADS-C was not direct in that the SADS-C includes only 25% of the ratings of the SADS-Part 1. Thus, the scales were formed on the basis of criterion items that remained on the 45 item measure. The single scale to undergo substantial revision was the Rare Symptoms Scale which included only three of the 10 items (refer to Appendix B for a listing of items by scale for the SADS-C). Similar to the SADS, it was anticipated that SADS-C classifications would not
match SIRS classifications but would produce results in
the expected direction.

**Development of Malingering Scales on the SADS and SADS-C**

The clinical interview continues to play a fundamental
role in diagnostic evaluation in mental health settings
(Rogers, 1996). According to Rogers, these interviews are
typically unstructured thereby providing clinicians with
the flexibility to adapt their interview to the needs of
the setting and/or the patient. This versatility makes
the formal evaluation of dissimulation impractical since
there is no standardized protocol from which to compare
data. In an attempt to address this problem, Rogers
(1996) has advocated the use of structured and semi-
structured interviews instead of unstructured clinical
interviews. Not only do these interviews cover a wide
range of symptomatology but they provide a systematized
method of recording data that includes severity ratings
and some observational data. These advantages provide
considerable control over examiner effects related to
style, breadth of information, and method of recording.

Research on the SADS and the SADS-C has identified
summary scales that assist in the assessment of
psychopathology (see, Endicott & Spitzer, 1979; Johnson,
Magaro, & Stern, 1986). However, neither of these
measures have been evaluated with regard to malingering.
The identification of malingering scales on the SADS and
the SADS-C would provide clinicians with a procedure for evaluating the validity of responses after collecting diagnostic data in a systematic manner.

**Screening for Mental Illness in a Jail Setting**

Lastly, a screening battery for malingering and mental illness was developed on the basis of the current and previous research findings on the RDS, SPS, SADS-C, and the M Test. The use of a diagnostic screening battery is considered advantageous in a jail setting since the measures can be administered quickly, require minimal training, and provide a breadth of information regarding need for mental health services. In addition, the battery used in the current study included measures of suicide risk and malingering which are two important areas of assessment in jail settings.

To summarize, the current study was designed to address the following five research questions:

1. How does dissimulation effect diagnostic screening instruments?
2. Is inconsistent responding a discriminating variable in the assessment of malingering?
3. Is there convergent validity between the SIRS and the M Test, SADS and SADS-C?
4. Is it possible to development of malingering scales on the SADS and SADS-C?
5. Is it feasible to develop a mental health screening battery for use in jail settings?
CHAPTER II

METHOD

Participants

The participants were 122 males incarcerated at the Tarrant County Jail between May 1995 and September 1995. Application of the SIRS criteria resulted in 24 participants being classified as suspected malingerers, 64 as patients, and 34 as indeterminant. The participants' ages ranged from 16 to 81 years with a mean age of 32.75 years (SD = 11.04). Two cases were not included in the analyses as psychological condition at the time of testing rendered initial evaluation impossible (e.g., severe dementia, and active hallucinations and agitation).

The groups were not expected to differ on the basis of demographic characteristics. Chi square analyses of race, referral source, and judicial status indicated that the three groups did not differ with respect to these variables. One-way ANOVAS were conducted on age, education level, number of previous arrests, number of times in jail (including the current incarceration), and number of times in prison. The results indicated that the three groups did not differ with respect to age or education but they did differ in terms of previous number of arrests, previous
number of times in jail, and previous number of times in prison. Malingers had significantly more arrests and jail admissions than both the patient and indeterminant groups while the indeterminant group had significantly more prison admissions than the patient group (refer to Table 2).

**Materials**

The measures used in the current study were the following: (a) Suicide Probability Scale (SPS), (b) Referral Decision Scale (RDS), (c) M Test, (d) Schedule of Affective Disorders and Schizophrenia - Change Version (SADS-C), (e) Structured Interview of Reported Symptoms (SIRS), (f) Schedule of Affective Disorders and Schizophrenia - Part 1 (SADS), and (g) Personality Assessment Inventory (PAI).

The SPS, SADS-C, and RDS are brief screening measures for the evaluation of psychopathology. While there are numerous mental health screening instruments, these three were chosen on the basis of prior research in correctional settings (see Rogers et al., 1995; Teplin & Schwartz, 1987) and the absence of research in the area of malingered suicidality. The SIRS was chosen as the gold standard from which to assess the construct validity the M Test and the SADS as potential screens for malingering. As already described, the SIRS was considered on appropriate standard on the basis of its sound psychometric properties. A description of each measure is provided below.
Table 2

Demographic and Background Characteristics of the Sample

<table>
<thead>
<tr>
<th></th>
<th>Groups</th>
<th></th>
<th></th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient</td>
<td>Indeterminant</td>
<td>Suspected Malingerer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>M</td>
<td>32.89</td>
<td>32.44</td>
<td>33.08</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>(12.27)</td>
<td>(9.79)</td>
<td>(9.86)</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>M</td>
<td>11.75</td>
<td>11.09</td>
<td>10.62</td>
<td>1.55</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>(2.62)</td>
<td>(3.68)</td>
<td>(1.97)</td>
<td></td>
</tr>
<tr>
<td>Arrests</td>
<td>M</td>
<td>7.02&lt;sub&gt;a&lt;/sub&gt;</td>
<td>15.79&lt;sub&gt;a&lt;/sub&gt;</td>
<td>19.77&lt;sub&gt;b&lt;/sub&gt;</td>
<td>4.14</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>(9.85)</td>
<td>(24.29)</td>
<td>(31.77)</td>
<td></td>
</tr>
<tr>
<td>Jail Admissions</td>
<td>M</td>
<td>6.03&lt;sub&gt;a&lt;/sub&gt;</td>
<td>14.45&lt;sub&gt;a&lt;/sub&gt;</td>
<td>19.36&lt;sub&gt;b&lt;/sub&gt;</td>
<td>4.33</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>(9.32)</td>
<td>(24.28)</td>
<td>(31.96)</td>
<td></td>
</tr>
<tr>
<td>Prison Admissions</td>
<td>M</td>
<td>0.44&lt;sub&gt;a&lt;/sub&gt;</td>
<td>1.12&lt;sub&gt;b&lt;/sub&gt;</td>
<td>0.83&lt;sub&gt;ab&lt;/sub&gt;</td>
<td>3.00</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>(1.02)</td>
<td>(1.75)</td>
<td>(1.37)</td>
<td></td>
</tr>
</tbody>
</table>

Note: The Student-Newman Keuls procedure was applied as the range statistic with a significance level = .05. Outliers (i.e., > 4 SDs) for number of arrests and number of times in jail were removed for these calculations (n = 2). Different subscripts denote significant differences between groups.
The Suicide Probability Scale

The Suicide Probability Scale (SPS) is a "36 item self-report measure designed to assist the clinician in the evaluation of suicide risk in adolescents and adults" (Golding, 1985, p. 1500). The instrument consists of four scales and provides the clinician with a weighted total score. The four scales on the SPS include (a) the Hopelessness Scale (HP; 12 items), (b) the Suicide Ideation Scale (SI; 8 items), (c) the Negative Self-Evaluation Scale (NSE; 9 items), and (d) the Hostility Scale (HS; 7 items).

The SPS was developed using both theoretical and empirical strategies. Scale development began with the generation of over 200 test items based on current theoretical explanations of suicide. Ultimately, the list was reduced to 36 items based on the ability of each item to discriminate between individuals who have attempted suicide, psychiatric patients who have not attempted suicide, and "normals" who have not attempted suicide. Although psychosocial stressors and past history of suicidality are not included in the scoring procedure, the manual stresses that these factors be considered when making final decisions.

Presumptive risk level (PRL) is a key factor in the interpretation of the SPS (e.g., high, intermediate, and low). PRL refers to "a weighting system that enables the examiner to choose a method for deriving the Probability
Score that is best suited for the context and purpose of the evaluation" (Cull & Gill, 1988, p. 5). According to the authors, the high risk category is recommended when evaluating individuals in such settings as suicide prevention centers, crisis clinics, and psychiatric inpatient facilities. The intermediate risk category is recommended for use with a general outpatient psychiatric population or psychiatric inpatients who display no evidence of suicidal ideation or major depression. Lastly, the low risk category is intended for use with the general population where a low presumptive level of risk is postulated.

Interpretation of the SPS begins with a translation of raw scores into T-score equivalents. As with many other psychometric instruments, the T-scores are considered to be clinically significant at \( T \geq 70 \). However, slightly lower T-Scores should not be discounted. For example, T-scores of 60 are said to require careful clinical evaluation of suicide risk whereas T-scores of 70 are considered to provide strong evidence for instituting suicide precautions. Interpretation of low scores, on the other hand, has not been extensively investigated. Nevertheless, the authors note that T-scores of 40 or below should alert the clinician to the possibility that the individual has consciously or unconsciously minimized their actual suicide potential.
The second step in the interpretation of the SPS is classification of presumptive risk and identification of the associated Probability Scores. Each weighted raw score is translated into a T-score that is associated with a particular probability score. Probability score refers to "the statistical likelihood that an individual belongs in the population of lethal suicide attempters as evidenced by their responses on the SPS" (p. 14); these scores do not refer to the likelihood that an individual will attempt suicide (Cull & Gill, 1988).

Classification of suicide risk then involves the translation of probability scores into categorical ratings that are associated with various treatment options. These ratings are labeled "severe" (probability score 75-100), "moderate" (probability score 50-74), "mild" (probability score 25-49), and "subclinical" (probability score 0-24). The standard error of measurement (SEM) is 2.99. Examples of treatment recommendations for patients in the severe range include immediate hospitalization, one-to-one monitoring, and psychotropic medication. In comparison, recommendations for patients in the subclinical range include routine monitoring and supportive intervention.

Internal consistency of the total SPS, the HP Scale and the SI Scale are reported to be good with alphas of .93, .85, and .88 respectively. The internal consistency of both the HS and the NSE scales are reported to be substantially
lower with alphas of .78 and .58. Test-retest reliability for the entire scale is noted to be excellent ($r = .95$) over a 10 day period (Golding, 1985).

**The Referral Decision Scale**

The RDS is a 15-item instrument designed as a rapid screen for the presence of schizophrenia, depression, and bipolar disorder in correctional institutions (Teplin & Schwartz, 1989). The RDS was empirically derived using data obtained on the NIMH Diagnostic Interview Schedule (DIS; Robins, Helzer, Croughan, & Ratcliff, 1981) using a large sample of jail detainees (Teplin & Schwartz, 1989). Items that were relevant to the identification of schizophrenia and mood disorders were retained and formed three subscales (i.e., schizophrenia, bipolar disorder, and major depression). The results of the initial validation study suggested that the three subscales had excellent sensitivity (.88) and specificity (.99) averaged across the three disorders. The researchers cross validated these results using a prison population and obtained good sensitivity (.79) and excellent specificity (.99).

Hart, Roesch, Corrado, and Cox (1993) conducted the first evaluation of the interrater reliability and construct validity of the RDS. In general, the results suggested that the RDS had acceptable reliability and validity for screening purposes with mentally disordered jail detainees. However, the authors also stated that the utility of the RDS
is limited by its high rate of false positives (i.e., overpredicting psychopathology) and the narrow range of disorders evaluated.

A more recent evaluation of the convergent and discriminant validity of the RDS using a multitrait-multimethod matrix (Campbell & Fiske, 1959) suggested that the scale is not effective at discriminating between schizophrenia, bipolar disorder and depression (Rogers, Sewell, Ustad, Reinhardt, & Edwards, 1995). The three scales were found to be moderately intercorrelated (M r = .44) with the majority of scale intercorrelations exceeding the coefficients for convergent validity using scales of the PAI and the SADS-C. The authors suggested that the RDS has potential utility as a screen for general psychiatric impairment in correctional settings but its use as a screen for specific disorders was not supported.

Schedule of Affective Disorders and Schizophrenia (SADS)

The SADS is a comprehensive semi-structured interview developed for the purposes of reducing information and criterion variance in clinical evaluations (Endicott & Spitzer, 1978). The interview is organized into two major sections based on the patient's current functioning and the patient's lifetime functioning. Part 1 of the SADS is organized by two time periods: (a) the most severe period in the most recent episode (within the past year) and (b) the week prior to the interview. The SADS-Part 2 was
developed to provide a measure of the patient's experience of psychiatric disturbance over their lifetime. The SADS was designed to yield both descriptive and diagnostic information regarding 26 diagnoses as delineated by the Research Diagnostic Criteria (RDC). The introduction chapter provides a comprehensive discussion of this measure in the evaluation of malingering.

The Schedule of Affective Disorders and Schizophrenia - Change Version (SADS-C)

The SADS-C is a semi-structured interview developed by Spitzer and Endicott (1978b) for the assessment of key psychiatric symptomatology. The interview consists of 45 critical symptoms that are extracted from the SADS-Part 1 and the Global Assessment Scale rated for the current time period. At present, the SADS-C depression scale has demonstrated good discriminant validity when compared to other depression screens (Endicott, Cohen, Nee, Fleiss, & Sarantakos, 1981; Johnson, Magaro, & Stern, 1986) and has been found to be useful in the measurement of treatment response in patients with bipolar disorder (Freeman, Clothier, Pazzaglia, Sesem, & Swann, 1992).

Johnson, Magaro, and Stern (1986) utilized a contrasted groups design to evaluate the use of the SADS-C as a diagnostic screen and symptom severity measure. Their population consisted of a normal control group and four patient groups (i.e., patients diagnosed with either
unipolar depression, bipolar disorder [manic], non-paranoid schizophrenia, or paranoid schizophrenia). The authors derived three subscales for the SADS-C (i.e., depression, schizophrenia, and mania) and assessed convergent validity with other diagnostic screening instruments. The results indicated that the SADS-C is a reliable screening instrument that is capable of differentiating among diagnostic groups.

A recent study of convergent and discriminant validity (Rogers et al., 1995) demonstrated that the SADS-C was an effective screening measure for the assessment of psychiatric symptomatology within the jail population. Specifically, the SADS-C was found to have both excellent discriminant validity and convergent validity when compared against the Personality Assessment Instrument (PAI).

A more recent study comparing the RDS, the SADS-C and the PAI suggested that the SADS-C is a useful screen for common mental disorders (Rogers et al., 1995). In addition, the authors evaluated the utility of a four-item subscale for anxiety that had not previously been defined. All of the subscales of the SADS-C displayed satisfactory convergent validity and excellent discriminant validity.

The M Test

The M Test is a 33-item, self-report inventory designed as a screen for malingered schizophrenia (Beaber et al., 1985). The inventory consists of true-false items that include true symptoms of schizophrenia, bizarre attitudes
and beliefs that are not symptoms of mental illness, and false symptoms suggestive of mental illness. In its original form, the measure was separated and scored on three sub-scales: the Schizophrenia Scale (10 items), the Confusion Scale (8 items), and the Malingering Scale (15 items). A recent revision of the scoring procedure (Rogers, Bagby, & Gillis, 1992) resulted in the formation of two, 10 item scales labeled the Rule-In and Rule-Out scales. The Rule-Out scale was designed to eliminate bona fide patients while the Rule-In scale was designed to identify potential malingerers. Like the SADS, the M Test is discussed in more detail in the introduction section.

The Structured Interview of Reported Symptoms

The SIRS is a structured interview designed to assess response styles with a particular emphasis on the feigning of mental disorders (Rogers, Bagby, & Dickens, 1992). The interview is composed of 172 items with eight primary and five supplementary scales. Items are scored on a three-point scale: A score of zero indicates a non-deviant response and scores of one and two reflect increasing levels of symptom endorsement. Classification of response style involves three categories (i.e., honest responding, indeterminate, and feigning) and is determined using either single scale scores or multiple scale scores. A more comprehensive discussion of the SIRS is provided in the introduction section.
Procedure

Setting

Participants were referred to the project by the Texas Department of Mental Health and Mental Retardation located at the Tarrant County Jail (TXMHMR-TCJ), Fort Worth, Texas. At this facility, TXMHMR-TCJ is the sole mental health service provider for inmates in the 10,000 bed jail facility. The services by TXMHMR-TCJ include evaluations for (a) psychiatric services (e.g., medication), (b) suicide risk, (c) dangerousness, (d) placement on specialized mental health units, and (e) outpatient placement upon release. On average, the TXMHMR-TCJ receives 800 referrals per month. These evaluations are typically initiated by one of three procedures: (a) a request by the screening officer at booking; (b) a request by the inmate at anytime during his or her incarceration; or (c) a request by a correctional officer when mental health problems are suspected.

Participant Selection

Each Monday morning twenty participants were randomly selected from the referral pool and twelve of these individuals were evaluated as part of the study. Over-selection of participants was necessary to ensure meeting the desired number of 12 participants per week. Participant attrition was due to release from the institution, non-availability of the inmate (e.g., meeting with lawyer, court date, or illness), or institutional sanctions that made
interviewing impossible. Individuals who were temporarily unavailable were interviewed the following week.

Administration of Measures

Evaluations were conducted on an individual basis, over two sessions by two examiners. Evaluations followed a standard procedure, were completed within two to four days and lasted for approximately 3.5 hours. The participants were approached by a correctional officer who informed them that a representative of the TXMHR-TCJ was here to conduct an evaluation. The participant would then be brought to the examiner, or vice versa depending on the unit, and an explanation of the purpose of the evaluation was provided. Specifically, the detainee was informed that (a) their name had been forwarded to the department for the purposes of conducting an initial mental health evaluation, and (b) their participation was voluntary and they could end the session at any time. Of the 122 participants, 82 completed the entire evaluation process, 20 completed the initial session and the SADS-Part 1, and 20 completed the initial session only.

The first stage of the evaluation involved the administration of the following instruments in the order presented: The RDS, SPS, M Test, SADS-C, and the SIRS. The SIRS was the last test administered in the first stage and was considered necessary to minimize bias related to examiner expectancy effects. Specifically, administration
of the SIRS may have provided clues to malingering status which could inadvertently have influenced the ratings of the SADS-C or the RDS. In addition, the results of the initial evaluation were masked in an effort to control experimenter expectancy effects at the second stage.

The second stage of the evaluation involved the administration of the SADS-Part 1 followed by the PAI. The mean number of days between first and second stages was 2.24, SD = 1.09). The order was standardized and was determined by the goals of the study. Since the SADS was of primary importance to the study, it was administered first to maximize the likelihood of participants completing this protocol. It was assumed that some participants would choose not to complete the evaluation process due to fatigue or possible restrictions imposed by the unit.

The order of tests was chosen to help maximize the likelihood that the majority of tests would be completed over both test sessions. Thus, after controlling for the effects of the SIRS, the decision was to order the tests by length of administration with the shorter tests administered first.

Interviewers

The interviews were conducted by three advanced graduate students in clinical psychology. Each interviewer had received didactic and practical training in the administration of diagnostic and structured interviews prior
to participation in this project. In addition, the interviewers had prior clinical experience with inmate populations through research conducted at a variety of correctional facilities including the Tarrant County Jail.

As already mentioned, the first stage of the evaluation was conducted by one of two interviewers, each conducted 61 of the 122 evaluations. The second stage of the evaluation was conducted by the primary investigator (N = 122).
CHAPTER III

RESULTS

Reliability of the Measures

Each measure was evaluated for internal consistency and item homogeneity. These analyses were conducted to ensure that each instrument met acceptable standards of reliability. Results obtained with unreliable measures can not be replicated or interpreted since large measurement error yields unstable results (DeVillis, 1991). Overall, the scales were found to be within acceptable limits for both reliability and homogeneity (refer to Table 3).

Malingering on Diagnostic Screening Instruments

The first research question was "How does dissimulation affect diagnostic screening instruments." This question was addressed using three measures designed for the rapid assessment of psychopathology: the RDS, SADS-C, and SPS. Tables 4, 5, and 6 provide means, standard deviations, and the results of multiple ANOVAS displaying group differences on each of the three symptom measures.

The results indicated that the three screening measures were susceptible to feigned responding. As evident in Table 4, 5, and 6 the mean scores for the suspected
Table 3

Reliability of Psychodiagnostic and Screening Instruments

<table>
<thead>
<tr>
<th></th>
<th>Alpha</th>
<th>Mean r</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RDS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>.83</td>
<td>.50</td>
</tr>
<tr>
<td>Depression</td>
<td>.62</td>
<td>.24</td>
</tr>
<tr>
<td>Bipolar</td>
<td>.59</td>
<td>.22</td>
</tr>
<tr>
<td><strong>SADS-C</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paranoia</td>
<td>.84</td>
<td>.63</td>
</tr>
<tr>
<td>Depression</td>
<td>.85</td>
<td>.39</td>
</tr>
<tr>
<td>Anxiety</td>
<td>.76</td>
<td>.35</td>
</tr>
<tr>
<td>Mania</td>
<td>.79</td>
<td>.42</td>
</tr>
<tr>
<td><strong>SPS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hopelessness</td>
<td>.85</td>
<td>.32</td>
</tr>
<tr>
<td>Suicide</td>
<td>.89</td>
<td>.49</td>
</tr>
<tr>
<td>Negative Self Evaluation</td>
<td>.69</td>
<td>.20</td>
</tr>
<tr>
<td>Hostility</td>
<td>.69</td>
<td>.24</td>
</tr>
<tr>
<td>SPS Total</td>
<td>.90</td>
<td>.20</td>
</tr>
<tr>
<td><strong>SIRS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rare Symptoms</td>
<td>.74</td>
<td>.26</td>
</tr>
<tr>
<td>Improbable &amp; Absurd Symptoms</td>
<td>.76</td>
<td>.31</td>
</tr>
<tr>
<td>Symptom Combinations</td>
<td>.72</td>
<td>.20</td>
</tr>
</tbody>
</table>

(Table continues)
Table 3 continued

<table>
<thead>
<tr>
<th>Alpha Mean  r</th>
<th>Mean r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blatant Symptoms</td>
<td>.89</td>
</tr>
<tr>
<td>Subtle Symptoms</td>
<td>.90</td>
</tr>
<tr>
<td>Reported vs. Observed</td>
<td>.75</td>
</tr>
</tbody>
</table>

M Test

<table>
<thead>
<tr>
<th></th>
<th>Mean r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule-In</td>
<td>.76</td>
</tr>
<tr>
<td>Rule-Out</td>
<td>.79</td>
</tr>
<tr>
<td>Total</td>
<td>.86</td>
</tr>
</tbody>
</table>

**Note:** Mean r = Average inter-item correlation. Scales SEV and SEL of the SIRS are not included since they are derived by arithmetic summing for which measures of consistency are not appropriate.

The malingerer group were significantly higher than those of the patient group and met the criteria for further psychiatric evaluation on all three measures.

Interestingly, the average score for the patient group was slightly below the recommended cut scores based on Teplin and Schwartz's data (1987). One possible explanation for this is that, on average, the patient group was not suffering from a severe mental illness that required treatment. An alternative hypothesis is that the RDS did
Table 4

Means, Standard Deviations and Significance Levels for the RDS by Group

<table>
<thead>
<tr>
<th>Groups</th>
<th>Patient Mean</th>
<th>Suspected Malingerer Mean</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schizophrenia</td>
<td>0.70</td>
<td>3.29</td>
<td>64.34</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>(1.18)</td>
<td>(1.73)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>1.58</td>
<td>3.25</td>
<td>27.58</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>(1.34)</td>
<td>(1.29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar</td>
<td>1.42</td>
<td>3.33</td>
<td>40.64</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>(1.30)</td>
<td>(1.13)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* The RDS was scored for endorsement of symptoms: scores of 2 on the schizophrenia subscale, 2 on the depression subscale, or 3 on the mania subscale suggest the need for further psychiatric evaluation.

Not represent the extent nor the range of psychopathology relevant to treatment needs. Some support for the latter hypothesis is found in the current study. Specifically, using the RDS as a single screen, 26 individuals were excluded from further evaluation. Inspection of the SPS and the SADS-C scores suggested that many of these individuals may have been experiencing significant symptomatology.
Table 5
Means, Standard Deviations, and Significance Levels for the SADS-C by Group

<table>
<thead>
<tr>
<th>SADS-C</th>
<th>Patient</th>
<th>Suspected Malingering</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>28.72</td>
<td>52.04</td>
<td>86.24</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>(8.24)</td>
<td>(15.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paranoia</td>
<td>8.28</td>
<td>13.17</td>
<td>37.39</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>(2.71)</td>
<td>(4.64)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>9.39</td>
<td>17.58</td>
<td>72.53</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>(3.23)</td>
<td>(5.64)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mania</td>
<td>5.70</td>
<td>6.92</td>
<td>3.25</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>(1.71)</td>
<td>(4.64)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Clinical range on the SADS-C was defined as follows: ≥ 34 on the depression scale, ≥ 10 on the mania scale, and ≥ 9 on the paranoia scale.

Specifically, six individuals scored in the clinical range on the SADS-C and 16 met criteria for further evaluation of suicidality (T-score of > 60 or < 40 on the SPS).

Consistency Between the SADS and the SADS-C

The second research question was "Is inconsistent responding a discriminating variable in the assessment of
Table 6

Means, Standard Deviations, and Significance Levels for the SPS by Group

<table>
<thead>
<tr>
<th>Groups</th>
<th>Patient M</th>
<th>SD</th>
<th>t</th>
<th>Suspected Malingierer M</th>
<th>SD</th>
<th>t</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP</td>
<td>19.32</td>
<td>7.36</td>
<td>61</td>
<td>32.71</td>
<td>10.74</td>
<td>73</td>
<td>43.73</td>
<td>.001</td>
</tr>
<tr>
<td>SI</td>
<td>12.40</td>
<td>5.50</td>
<td>60</td>
<td>23.75</td>
<td>9.26</td>
<td>71</td>
<td>49.02</td>
<td>.001</td>
</tr>
<tr>
<td>NSE</td>
<td>15.19</td>
<td>4.80</td>
<td>59</td>
<td>19.42</td>
<td>4.21</td>
<td>67</td>
<td>14.29</td>
<td>.001</td>
</tr>
<tr>
<td>HS</td>
<td>12.66</td>
<td>4.95</td>
<td>63</td>
<td>18.29</td>
<td>6.29</td>
<td>71</td>
<td>19.18</td>
<td>.08</td>
</tr>
<tr>
<td>TS</td>
<td>59.58b</td>
<td>16.51</td>
<td>64</td>
<td>94.17a</td>
<td>22.59</td>
<td>74</td>
<td>61.27</td>
<td>.08</td>
</tr>
</tbody>
</table>

Note. t = t-Score; HP = Hopelessness Scale; SI = Suicidal Ideation Scale; NSE = Negative Self Evaluation Scale; HS = Hostility Scale; TS = Total Score. Superscripts denote presumptive risk level: a = severe risk; b = mild risk.
malingering"? The hypothesis that suspected malingerers would respond inconsistently over time was evaluated using multivariate analysis of variance (MANOVA) and correlational statistics. Identical items on the SADS and SADS-C were paired and then difference scores were calculated (e.g., the absolute difference between SADS item #235 and SADS-C item #213). Seventeen of the 27 comparisons showed significant differences with the patient group demonstrating more consistency than either the indeterminant or suspected malingerer groups (see Table 7). As expected, patient group demonstrated the most consistency, the suspected malingerer group the least, and the indeterminant was between the two.

A limitation of these ANOVAs was the restricted and skewed distribution in scores. A repeated measures MANOVA based on total rather than difference scores was contemplated. Although this would address partly the problems with restricted and skewed range, it would obscure the absolute differences across time periods. Therefore, to maximize the clinical relevance of the findings, ANOVAs were computed on difference scores.

To evaluate the effects of inconsistency on aggregate scores, the difference scores (as described previously) were summed and averaged across three subscales of the SADS-C. Frequency counts of absolute difference scores were tabulated for each group and reported in Table 8. In
Table 7

Means and Standard Deviations of Difference Scores on the
SADS and SAD-C by Group

<table>
<thead>
<tr>
<th>Item Pairs</th>
<th>Patient</th>
<th>Indeter.</th>
<th>Suspected Maling.</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>SADS-SADS-C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>235-213</td>
<td>.80</td>
<td>1.10</td>
<td>1.18</td>
<td>1.87</td>
<td>.16</td>
</tr>
<tr>
<td></td>
<td>.90</td>
<td>1.16</td>
<td>.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>239-214</td>
<td>.68</td>
<td>1.03</td>
<td>.86</td>
<td>1.57</td>
<td>.21</td>
</tr>
<tr>
<td></td>
<td>.87</td>
<td>.93</td>
<td>.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>241-215</td>
<td>.50&lt;sub&gt;a&lt;/sub&gt;</td>
<td>1.00&lt;sub&gt;b&lt;/sub&gt;</td>
<td>1.04&lt;sub&gt;ab&lt;/sub&gt;</td>
<td>4.04</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>.68</td>
<td>1.23</td>
<td>1.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>243-216</td>
<td>.60</td>
<td>1.13</td>
<td>.77</td>
<td>2.41</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td>1.05</td>
<td>1.17</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>245-217</td>
<td>.60</td>
<td>.77</td>
<td>.73</td>
<td>0.35</td>
<td>.71</td>
</tr>
<tr>
<td></td>
<td>.95</td>
<td>1.19</td>
<td>.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>247-218</td>
<td>.20&lt;sub&gt;a&lt;/sub&gt;</td>
<td>.73&lt;sub&gt;b&lt;/sub&gt;</td>
<td>1.05&lt;sub&gt;b&lt;/sub&gt;</td>
<td>9.41</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>.57&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.02&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>264-219</td>
<td>.64&lt;sub&gt;a&lt;/sub&gt;</td>
<td>1.17&lt;sub&gt;b&lt;/sub&gt;</td>
<td>1.27&lt;sub&gt;ab&lt;/sub&gt;</td>
<td>3.72</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>1.01</td>
<td>1.23</td>
<td>1.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>266-220</td>
<td>.72&lt;sub&gt;a&lt;/sub&gt;</td>
<td>1.33&lt;sub&gt;b&lt;/sub&gt;</td>
<td>1.23&lt;sub&gt;ab&lt;/sub&gt;</td>
<td>4.09</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>.95&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.16</td>
<td>1.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>268-221</td>
<td>.60</td>
<td>.90</td>
<td>1.05</td>
<td>1.68</td>
<td>.19</td>
</tr>
<tr>
<td></td>
<td>1.01</td>
<td>1.16</td>
<td>1.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>271-222</td>
<td>.22&lt;sub&gt;a&lt;/sub&gt;</td>
<td>1.03&lt;sub&gt;b&lt;/sub&gt;</td>
<td>1.82&lt;sub&gt;c&lt;/sub&gt;</td>
<td>22.18</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>.68</td>
<td>1.30</td>
<td>1.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>313-223</td>
<td>.88</td>
<td>1.10</td>
<td>1.09</td>
<td>0.59</td>
<td>.56</td>
</tr>
<tr>
<td></td>
<td>1.14</td>
<td>1.13</td>
<td>1.02</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Table continues)
Table 7 continued

<table>
<thead>
<tr>
<th>Item Pairs</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>316-227</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>318-228</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>323-229</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>327-230</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>331-231</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>333-232</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>341-233</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>348-234</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>354-235</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>356-236</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>358-237</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>360-238</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>362-239</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>419-240</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
</tbody>
</table>

(Table continues)
### Table 7 continued

<table>
<thead>
<tr>
<th>Item Pairs</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>450-242</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>529-244</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
</tbody>
</table>

**Note:** Column A = item number associated with the SADS-Part 1. Column B = item numbers associated with the SADS-C. Student-Newman Keuls procedure was applied as the range statistic (α = .05). Different subscripts denote statistically significant differences between groups. Patient = Patient Group. Suspected Maling. = Suspected Malingener Group. Indeter. = Indeterminant Group.

In addition, correlations of identical items were calculated and averaged by scale for each group. With the exception of the mania scale, the results were as hypothesized with the patient group responding more consistently than both comparison groups across two test periods.

**Convergent Validity Between the SIRS and the M Test, SADS and SADS-C**

The third research question was "Is there convergent validity between the SIRS and the M Test, SADS and SADS-C?"
Table 8

Consistent Responding by Group on the Scales of the SADS-C

<table>
<thead>
<tr>
<th>Subscales</th>
<th>Consistent</th>
<th>Inconsistent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt; 1</td>
<td>2</td>
</tr>
<tr>
<td><strong>Depression (15 items)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient</td>
<td>87.6%</td>
<td>7.9%</td>
</tr>
<tr>
<td>Indeterminant</td>
<td>72.8%</td>
<td>15.1%</td>
</tr>
<tr>
<td>Suspected Malingerer</td>
<td>70.0%</td>
<td>16.7%</td>
</tr>
<tr>
<td><strong>Anxiety (4 items)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient</td>
<td>90.0%</td>
<td>5.0%</td>
</tr>
<tr>
<td>Indeterminant</td>
<td>76.7%</td>
<td>10.0%</td>
</tr>
<tr>
<td>Suspected Malingerer</td>
<td>56.8%</td>
<td>13.6%</td>
</tr>
<tr>
<td><strong>Mania (5 items)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient</td>
<td>97.4%</td>
<td>0.8%</td>
</tr>
<tr>
<td>Indeterminant</td>
<td>90.6%</td>
<td>6.0%</td>
</tr>
<tr>
<td>Suspected Malingerer</td>
<td>79.1%</td>
<td>2.7%</td>
</tr>
<tr>
<td><strong>Other (5 items)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient</td>
<td>90.0%</td>
<td>4.8%</td>
</tr>
<tr>
<td>Indeterminant</td>
<td>72.0%</td>
<td>16.0%</td>
</tr>
<tr>
<td>Suspected Malingerer</td>
<td>65.4%</td>
<td>16.4%</td>
</tr>
</tbody>
</table>
This question was addressed in a variety of ways using correlational statistics, cross-tabulation, and step-wise discriminant function analysis. Comparisons were made between the patient group and the suspected malingerer group for each measure.

Convergent Validity of the M Test and the SIRS

The convergent validity between the M Test (Rule-In and Rule-Out scales) and the SIRS was evaluated using correlational statistics and cross-tabulation. As predicted, the scales of the M Test show high positive correlations with the eight clinical scales of the SIRS. The only exception was the Reported vs. Observed symptoms which, although significant and positive, correlated much less than the other scales (refer to Table 9).

As mentioned previously, Rogers, Bagby, and Gillis (1992) developed two options for use of the Rule-In/Rule-Out procedure: Option A (Rule-Out score of < 4 and Rule-In score of < 2) is considered to be more useful in general clinical practice and Option B (Rule-Out score of < 4 and a Rule-In score of = 0) is thought to be more useful in forensic settings. In the current study, Option A eliminated 82.5% of the patients (vs. 83.8% for the Rogers study) and retained 87% of the suspected malingerers (vs. 83.0% for the Rogers study) and Option B eliminated 69.8% of the patients (vs. 70.6% for the Rogers study) and retained
Table 9
Correlations Between the SIRS and the M Test

<table>
<thead>
<tr>
<th>SIRS Scales</th>
<th>Rule-In</th>
<th>Rule-Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rare Symptoms</td>
<td>.69**</td>
<td>.64**</td>
</tr>
<tr>
<td>Improbable and Absurd Symptoms</td>
<td>.73**</td>
<td>.54**</td>
</tr>
<tr>
<td>Symptom Combinations</td>
<td>.71**</td>
<td>.66**</td>
</tr>
<tr>
<td>Blatant Symptoms</td>
<td>.67**</td>
<td>.67**</td>
</tr>
<tr>
<td>Subtle Symptoms</td>
<td>.44**</td>
<td>.62**</td>
</tr>
<tr>
<td>Selectivity of Symptoms</td>
<td>.58**</td>
<td>.73**</td>
</tr>
<tr>
<td>Severity of Symptoms</td>
<td>.50**</td>
<td>.57**</td>
</tr>
<tr>
<td>Reported vs. Observed Symptoms</td>
<td>.34**</td>
<td>.36**</td>
</tr>
</tbody>
</table>

Note. For all rs, "p < .01.

95.7% of the suspected malingerers (vs. 95.2% for the Rogers study). Optimal cut scores for the current sample were the same as those of Rogers, Bagby, and Gillis (1992) and classification statistics were nearly identical to those of the original validation study. Tables 10 and 11 provide the classifications for each option.

Convergent Validity between the SADS and the SIRS

Convergent validity of the SADS models of malingering and the SIRS was evaluated with three methods:
Correlations, cross-tabulation with classification, and discriminant function analysis. First, three scales of the SADS (Rare Symptom, Overendorsement of Symptoms, and Symptom Severity) were correlated with three conceptually similar scales of the SIRS. Evidence for convergent validity was found for all three scales (refer to Table 12). As expected, the correlations between the two severity scales were the lowest of the planned comparisons as these scales are considered to be conceptually less similar than the other two scales. Unexpectedly, the SEV scale of the SIRS evidenced higher correlations with both RS and SEL than with the SEV scale of the SADS. This pattern was also evident for the RS scale of the SIRS, where the SEL was more highly correlated than RS. These intercorrelations suggest that the RS and SEL of the SADS assess similar constructs to all three of the SIRS scales.

The convergent validity of the SADS Threshold Model and Clinical Decision Model were evaluated using cross-tabulation and discriminant function analysis. Application of the Threshold Model to a sample \( (n = 72) \) of suspected malingerers and patients produced an overall hit rate of 70.8%. The model was very effective at classifying suspected malingerers with all 22 being accurately classified (sensitivity = 1.0). However, the classification rate for the patient group was poor and resulted in accurate classification at just above chance (specificity = .58).
Table 10

Predictive Accuracy of the M Test in the Detection of Malingering - Option A

### Step 1

<table>
<thead>
<tr>
<th>Predicted Group</th>
<th>Suspected Malingerer</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspected Malingerer</td>
<td>23</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>(45.1%)</td>
<td>(54.9%)</td>
</tr>
<tr>
<td>Patient</td>
<td>0</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>(0.0%)</td>
<td>(100.0%)</td>
</tr>
</tbody>
</table>

### Step 2

<table>
<thead>
<tr>
<th>Predicted Group</th>
<th>Suspected Malingerer</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspected Malingerer</td>
<td>20</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>(64.5%)</td>
<td>(35.5%)</td>
</tr>
<tr>
<td>Patient</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>(15.0%)</td>
<td>(85.0%)</td>
</tr>
</tbody>
</table>

Hit Rate = 83.7%
Sensitivity = .87
Specificity = .83
Positive Predictive Power = .63
Negative Predictive Power = .95
Table 11

Predictive Accuracy of the M Test in the Detection of Malingering - Option B

Step 1

<table>
<thead>
<tr>
<th>Predicted Group</th>
<th>Suspected Malingerer</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspected Malingerer</td>
<td>23 (45.1%)</td>
<td>28 (54.9%)</td>
</tr>
<tr>
<td>Patient</td>
<td>0 (0.0%)</td>
<td>35 (100.0%)</td>
</tr>
</tbody>
</table>

Step 2

<table>
<thead>
<tr>
<th>Predicted Group</th>
<th>Suspected Malingerer</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspected Malingerer</td>
<td>22 (53.7%)</td>
<td>19 (46.3%)</td>
</tr>
<tr>
<td>Patient</td>
<td>1 (10.0%)</td>
<td>9 (90.0%)</td>
</tr>
</tbody>
</table>

Hit Rate = 76.7%
Sensitivity = .96
Specificity = .70
Positive Predictive Power = .54
Negative Predictive Power = .98
Table 12

Convergent Validity Between the SADS-Part 1 and the SIRS

<table>
<thead>
<tr>
<th>SIRS Scales</th>
<th>RS</th>
<th>SEL</th>
<th>SEV</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEV</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. SADS Scales: RS = Rare Symptoms; SEL = clinically endorsed symptoms ($\geq 3$ and $< 5$); SEV = symptoms reported at a severe level ($\geq 5$). SIRS Scales: RS = Rare Symptoms; SEL = Selectivity of Symptoms; SEV = Severity of Symptoms.

PPP for this model was low at .51 while NPP was excellent at 1.0.

The overall hit rate for the Clinical Decision Model, using the same sample, was substantially higher at 81.9%. However, only 10 of 22 malingerers were accurately classified (sensitivity = .46). Improvement in classification was the result of increased accuracy with respect to the patient group (specificity = .96). Both PPP and NPP predictive power were moderately high at .83 and .80 respectively.
Four of the five criteria of the decision models were entered into a step-wise discriminant function analysis to determine which scales have the best predictive validity (refer to Table 13). All four of the scales remained in the function and together accounted for 51.8% of the variance in classification.

In response to a recent observation that rare symptoms do not adequately discriminate between suspected malingerers and patients in acute psychotic states (Rogers, in press) the discriminant analysis was repeated without the Rare Symptom Scale as a predictor variable. The removal of this variable resulted in a decrease in all of the classification statistics suggesting that it is a useful predictor variable in this sample (hit rate = 83.3%; sensitivity = .68, specificity = .90, positive predictive power = .75, negative predictive power = .87). All three scales remained in the function.

Convergent Validity between the SADS-C and the STIRs

As mentioned previously, the items included in the SADS-C Threshold Model and Clinical Decision Model were those that are identical or conceptually similar in content to those already developed for the SADS.

The overall hit rate for the SADS-C Threshold Model was 81.8% (n = 88) and sensitivity and specificity were high at .88 and .80 respectively. PPP for the model was moderate at
### Table 13

**Discriminant Analysis for the SADS Decision Model**

<table>
<thead>
<tr>
<th>Predicted Group</th>
<th>Actual Group</th>
<th>Suspected</th>
<th>Malingerer</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspected</td>
<td>17</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malingerer</td>
<td>(94.4%)</td>
<td>(5.6%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient</td>
<td>5</td>
<td>49</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(9.3%)</td>
<td>(90.7%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Wilks' lambda = .49

$\chi^2 (4, n = 72) = 48.78, \ p < .001$

Canonical correlation = .72

Hit Rate = 91.7%

Sensitivity = .77

Specificity = .98

Positive Predictive Power = .94

Negative Predictive Power = .91

**Note.** The scales entered in the discriminant function were: Rare Symptoms, Contradictory Symptoms, Unusual Symptom Combinations, and Symptom Severity ($\geq 5$).
.62 while NPP was high at .94. The Clinical Decision Model produced a slightly higher hit rate (84.1%) with low sensitivity (.46) and excellent specificity (.98). Both PPP and NPP were high at .92 and .83 respectively.

Once again, four of the five criteria\(^7\) of the decision models were entered into a discriminant function analyses to determine which scales have the best predictive validity\(^8\) (refer to Table 14).

Once again, the utility of rare symptoms was evaluated using a discriminant function that excluded the RS scale leaving three of the four predictor variables in the equation (contradictory symptoms, unusual symptom combinations and number of items endorsed \(\geq 3\)). Ultimately, the discriminant function using suspected malingerers and patients as dependent variables included two of the three predictor variables: Unusual symptom combinations and number of items scored \(\geq 3\). All of the classification statistics remained almost identical with the removal of rare symptoms and concomitant removal of contradictory symptom pairs (hit rate = 90.8\%, sensitivity = .88, specificity = .92, PPP = .81, NPP = .95).\(^9\)

**Identification of Malingering Scales on the SADS and the SADS-C**

The fourth research question was "Is it possible to development of malingering scales on the SADS and SADS-C?"
Table 14

**Discriminant Function Analysis for the SADS-C Decision Model**

<table>
<thead>
<tr>
<th>Predicted Group</th>
<th>Actual Group</th>
<th>Suspected</th>
<th>Malingerer</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspected</td>
<td></td>
<td>21</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Malingerer</td>
<td></td>
<td>(84.0%)</td>
<td>(16.0%)</td>
<td></td>
</tr>
<tr>
<td>Patient</td>
<td></td>
<td>3</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4.8%)</td>
<td>(95.2%)</td>
<td></td>
</tr>
</tbody>
</table>

Wilks' lambda = .43

χ² (4, n = 88) = 70.53, p < .001

Canonical correlation = .75

Hit Rate = 92.0%

Sensitivity = .88

Specificity = .94

Positive Predictive Power = .84

Negative Predictive Power = .95

**Note.** Rare Symptom, Contradictory Symptom Pairs, Unusual Symptom Combinations, and Overendorsement of Symptoms (≥ 3) were included in the equation (in the order above).
This question was addressed using a combination of Chi Square analysis, ANOVA, and step-wise discriminant function analysis. Scale development is outlined in the following paragraphs and is described separately for each measure.

Scale Development for the SADS

The procedure for scale development was exploratory in design involving 181 of the rated items on the SADS-Part 1 (153 items scored for the worst period in the current episode and 28 items scored for the past week). In order to reduce the number of variables to a manageable number, a series of ANOVAs (for interval data) and Chi square analyses (for nominal data) were conducted. Of these 181 analyses, 20 comparisons showed significant differences ($p < .001$) between the suspected malingerers ($n = 22$) and the patient ($n = 50$) groups. These variables were then entered into a stepwise discriminant function analysis for the purpose of identifying the most useful discriminating variables and eliminating those that were unnecessary (Klecka, 1980). The analysis produced a Wilks' lambda of $.27$, $\chi^2 (9, n = 72) = 84.00$ ($p < .001$). The canonical correlation was $.85$, indicating that these criteria accounted for 72.3% of the variance in explaining feigned versus nonfeigned self report. The overall classification rate was excellent at 91.7% with accurate classification of 81.3% of the suspected malingering group (sensitivity = .81; PPP = .90) and 96.0% of the patient group (specificity = .96; NPP = .92). Nine
of the 20 variables were retained in the discriminant function and were included on the resulting scale (refer to Appendix C for an item description).

Two scoring procedures were applied to the nine-item scale: (a) sum of the scores on each item (SADS-Sum) and (b) a count of the items that were endorsed at ≥ 3 (items considered to be clinically significant; SADS-Count). Reliability and item homogeneity were evaluated for SADS-Sum. Analysis indicated that SADS-Sum exhibits good internal consistency with a Cronbach's coefficient alpha of .81, and an acceptable level of homogeneity (.32) (Briggs & Cheek, 1986).

The results of two ANOVAs revealed that the suspected malingerers obtained significantly higher scores than patients on both SADS-Sum and SADS-Count (refer to Table 15). Optimum cut scores were identified for both scales using the cross-tabulation procedure; refer to Tables 16 and 17 for classification values. For the SADS-Count, decreasing in the cut score to 1 improved the classification of suspected malingerers to 90.9%. However, the overall hit rate decreased to 84.7% with similar decreases in specificity (.82) and positive predictive power (.69). Negative predictive power remained high at .95. A cut score of 1 is considered more appropriate when the identification of suspected malingerers is considered most paramount.
Scale Development for the SADS-C

Development of a malingering scale on the SADS-C followed the same exploratory procedures as the SADS-Part 1. Item reduction began with 36 ANOVAs and 3 Chi Square analyses resulting in identification of 13 items that differentiated suspected malingerers (n = 24) from patients (n = 64) (p < .001). These variables were then entered into a step-wise discriminant function analysis which produced a Wilks' lambda of .37 and $\chi^2 (6, n = 88) = 81.76$ (p < .001). The canonical correlation was .79, indicating that these criteria accounted for 62.4% of the variance in explaining feigned versus nonfeigned self report. The overall classification rate was excellent at 92.1% with very good sensitivity (.83) and superb specificity (.95). Positive and negative predictive power were high at .87 and .94 respectively. Six of the 13 variables were retained in the discriminant function and were included on the scale (refer to Appendix D for an item description).

The resulting scale was then scored in the same manner as the SADS-Part 1: (a) sum of the values on the items (SADS-C-Sum) and (b) a count of the items that were endorsed at $\geq$ 3 (SADS-C-Count). Scale reliability and item homogeneity for the SADS-C-Sum were satisfactory with alpha of .79, and level of homogeneity equal to .38.
Table 15
Means, Standard Deviations, and Significance Levels for the SADS by Group

<table>
<thead>
<tr>
<th>Groups</th>
<th>Patient M</th>
<th>Suspected Malingger M</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>SADS-Sum</td>
<td>12.08</td>
<td>24.09</td>
<td>103.16</td>
<td>.001</td>
</tr>
<tr>
<td>SD (3.24)</td>
<td>(6.84)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SADS-Count</td>
<td>0.98</td>
<td>5.45</td>
<td>104.74</td>
<td>.001</td>
</tr>
<tr>
<td>SD (1.30)</td>
<td>(2.40)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. SADS-Sum = score reflects the sum of a nine-item scale of continuous variables; SADS-Count = score reflects the sum of a nine-item scale using a dichotomous scoring procedure (Scoring: 0 for values < 3 and 1 for values ≥ 3).

The results of two ANOVAs revealed that suspected malingerers obtained significantly higher scores on both SADS-C-Sum and SADS-C-Count than did the patients (refer to Table 18).

Two cut scores were identified for the SADS-C-Sum. Option 1 is a cut score of 11 and Option 2 is cut score of 15. Suspected malingerers typically score above these cut scores. In accordance with Rogers's categorizations, Option 1 is termed the Threshold Model and is recommended for use
Table 16

Predictive Accuracy of the SADS-Sum by Group

<table>
<thead>
<tr>
<th>Predicted Group</th>
<th>Suspected</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Malingerer</td>
<td></td>
</tr>
<tr>
<td>Suspected</td>
<td>22</td>
<td>6</td>
</tr>
<tr>
<td>Malingering</td>
<td>(78.6%)</td>
<td>(21.4%)</td>
</tr>
<tr>
<td>Patient</td>
<td>2</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>(4.3%)</td>
<td>(95.7%)</td>
</tr>
</tbody>
</table>

Hit Rate = 88.8%
Sensitivity = .91
Specificity = .88
Positive Predictive Power = .79
Negative Predictive Power = .96

Note. SADS-Sum = score reflects the sum of a nine-item scale of continuous variables. Cut score > 14 for malingering. Highest possible score = 44.
Table 17

Predictive Accuracy of the SADS-Count by Group

<table>
<thead>
<tr>
<th>Predicted Group</th>
<th>Suspected Malingerer</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspected Malingerer</td>
<td>18 (78.3%)</td>
<td>5 (21.7%)</td>
</tr>
<tr>
<td>Patient</td>
<td>4 (8.16%)</td>
<td>45 (91.8%)</td>
</tr>
</tbody>
</table>

Hit Rate = 87.5%
Sensitivity = .82
Specificity = .90
Positive Predictive Power = .78
Negative Predictive Power = .92

Note. SADS-Count = score reflects the count of a nine-item scale using a dichotomous scoring procedure (Scoring: 0 for values < 3 and 1 for values ≥ 3). Cut score > 2 for malingering. Highest possible score = 9.
Table 18
ANOVA: Patient versus Suspected Malingering Groups on the SADS-C-Sum and SADS-C-Count

<table>
<thead>
<tr>
<th>Groups</th>
<th>Suspected</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient</td>
<td>Malingerer</td>
<td>F</td>
<td>p</td>
</tr>
<tr>
<td>SADS-C-Sum</td>
<td>8.58</td>
<td>19.17</td>
<td>135.19</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>(8.58)</td>
<td>(6.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SADS-C-Count</td>
<td>0.70</td>
<td>3.83</td>
<td>137.85</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>(0.87)</td>
<td>(1.61)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. SADS-C-Sum = score reflects the sum of a six-item scale of continuous variables; SADS-C-Count = score reflects the sum of a six-item scale using a dichotomous scoring procedure (0 = < 3; 1 = ≥ 3).

as a very general screen when sensitivity and NPP is of greatest concern. Option 2 is termed the Clinical Decision Model and is recommended when high PPP and reduction of false positives is desired (see Tables 19 and 20). These scales have only undergone initial validation and thus, are currently recommended for research purposes only.
Table 19

Predictive Accuracy of the SADS-C-Sum by Group - Threshold Model

Option 1 (cut score > 11)

<table>
<thead>
<tr>
<th>Predicted Group</th>
<th>Suspected</th>
<th>Malingerer</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspected</td>
<td>22</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Malingerer</td>
<td>22 (75.9%)</td>
<td>7 (24.1%)</td>
<td></td>
</tr>
<tr>
<td>Patient</td>
<td>2</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (3.4%)</td>
<td>57 (96.6%)</td>
<td></td>
</tr>
</tbody>
</table>

Hit Rate = 89.8%
Sensitivity = .92
Specificity = .89
Positive Predictive Power = .76
Negative Predictive Power = .97

The optimum cut score for SADS-C-Count was ≥ 3 with suspected malingerers typically scoring three or greater. This cut score produced a hit rate of 92.1% with 95.3% of the patients and 83.3% of the suspected malingerers (specificity = .95, sensitivity = .83). PPP and NPP were
Table 20

**Classification Accuracy of the SADS-C-Sum by Group - Clinical Decision Model**

<table>
<thead>
<tr>
<th>Predicted Group</th>
<th>Suspected</th>
<th>Malingerer</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspected</td>
<td>19</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Malingerer</td>
<td>(95.0%)</td>
<td>(5.0%)</td>
<td></td>
</tr>
<tr>
<td>Patient</td>
<td>5</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.4%)</td>
<td>(92.6%)</td>
<td></td>
</tr>
</tbody>
</table>

**Hit Rate = 93.2%**

**Sensitivity = .79**

**Specificity = .98**

**Positive Predictive Power = .95**

**Negative Predictive Power = .93**

Hit rate remained high at .87 and .94 respectively. Decreasing the cut score to ≥ 2 improved the classification of suspected malingerers to 91.7% (20/22) but decreased both the hit rate (85.2%) and the accuracy of classification of patients (82.8%).
Development of a Screening Battery for Malingering and Mental Illness

The fifth research hypothesis was "Is it feasible to develop a mental health screening battery for use in jail settings? The RDS, SPS, SADS-C, and M Test were combined to form a three-stage interpretive strategy for the screening of malingering and mental illness. Only those participants who completed all of the screening measures (n = 84) were included in the analyses.

Stage 1 - Do the reported symptoms require further evaluation?

Stage 1 of the procedure included basic screening measures, the RDS and the SPS, that were used to identify individuals whose symptoms were sufficient to warrant further evaluation. Individuals that did not meet criteria for suspected mental illness (i.e., none of the RDS criteria were met) or suicide risk (T-scores between 40 and 60 on the SPS) were excluded from further evaluation.

Stage 2 - Are the symptoms indicative of a major mental disorder?

Stage 2 involved the interpretation of the SADS-C criteria for mental illness and malingering. Individuals who scored below the clinical range on the SADS-C were excluded from further evaluation\(^\text{10}\). Subsequently, criteria for suspected malingering on the SADS-C-Sum was applied to the remaining sample. In order to retain the majority of
malingers, the more liberal cut score (> 11) on the SADS-C-Sum was used. The more liberal cut score resulted in more patients and suspected malingerers being retained for further evaluation. As previously mentioned, this cut score represents the "Threshold Model" in which sensitivity and negative predictive power are of primary importance.

Stage 3 - Are the symptoms genuine?

Stage 3 involved interpretation of the M Test using the Rule-In/Rule-Out procedure with Option B criteria. Once again, the choice to use the "Threshold Model" was based on the decision to maximize sensitivity such that the majority of suspected malingerers would be retained for further evaluation. Scores of > 4 on the Rule-Out scale resulted in further evaluation of malingering using the Rule-In scale. Score of > 4 on the Rule-Out Scale and > 0 on the Rule-in scale suggested the need for further evaluation of mental illness and malingering. Figure 1 outlines the procedure, the classification rates, and the recommendations for further evaluation.¹¹

The screening process demonstrated superb classificatory accuracy. The hit rate was excellent at 90.9% with equally high sensitivity (.91) and specificity (.97). PPP and NPP were .91 and .97 respectively.
**Figure 1**

**Screening Battery for Malingering and Mental Illness**

<table>
<thead>
<tr>
<th>RETAIN</th>
<th>SCREEN OUT (recommendations)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>84</strong> (SAMPLE)</td>
<td></td>
</tr>
<tr>
<td>Suspected Malingering</td>
<td>Patient</td>
</tr>
<tr>
<td><strong>Stage 1</strong></td>
<td></td>
</tr>
<tr>
<td>RDS &amp; SPS</td>
<td>RDS=0 8 No need for further evaluation</td>
</tr>
<tr>
<td>SPS T Between 40 &amp; 60</td>
<td>1</td>
</tr>
<tr>
<td><strong>Stage 2</strong></td>
<td></td>
</tr>
<tr>
<td>SADS-C</td>
<td>SADS-C=0 27 No need for further evaluation</td>
</tr>
<tr>
<td><strong>Stage 3</strong></td>
<td></td>
</tr>
<tr>
<td>M Test</td>
<td>RO &lt; 4 4 Evaluate for MI only</td>
</tr>
<tr>
<td></td>
<td>RO &lt; 4 &amp; RI=0 1 Evaluate for MI only</td>
</tr>
</tbody>
</table>

Evaluate for Malingering
CHAPTER IV

DISCUSSION

Correctional settings pose unique challenges for the assessment of malingering and other response styles. The need for rapid identification of mentally ill detainees at times supersedes the need for comprehensive evaluations that include assessment of response style. Unfortunately, such a compromise can lead to misallocate resources and inadvertently effect future judicial proceedings. As an example, a detainee may feign mental disorder in a jail facility in hopes that it would substantiate a future plea of incompetence or not criminally responsible. What is clear is the need for an expedient and cost effective screening of malingering. However, the question remains as to what methodology best addresses the need for efficiency while maintaining adequate positive predictive power.

Past research in correctional settings typically have addressed the efficacy of single measures as gross screens of current symptomatology. In contrast, the current study provides a two-level assessment: (a) mental health screening and (b) structured clinical interviewing. This study addressed three relevant clinical issues. First, the results show that screening instruments are easily feigned.
Second, the study provides preliminary data on a screening battery that may be useful in identifying potential feigners and patients with severe symptoms that require further evaluation. Third and most importantly, it provides empirical validation for the SADS models of malingering for use with two versions: the SADS and the abbreviated SADS-C. Taken together, the results suggest that general mental health evaluations are susceptible to feigning and that early detection is possible.

Two features of the methodology increased the study's external validity. The first is the use of known-groups research comparison. This design is especially useful because suspected malingerers are compared to genuine patients. Moreover, the use of contrasted groups from a correctional setting further enhances its generalizability. The more commonly used research method is the simulation design. This latter design only permits the study of how compliant participants produce feigned presentations; these results are then extrapolated to clinical and correctional settings.

The second design improvement was the use of a standardized, psychometric measure as the gold standard for malingering. Alternative standards include the use of clinical judgements from forensic experts or a decision rule based on a combination of tests. Although no ideal method is accepted to operationally define malingering, the current
study used the SIRS, an extensively validated measure that provides a high level of accuracy in the evaluation of malingering. The SIRS was considered an appropriate standard since its sound psychometric properties and structured administration provide an identifiable standard from which to validate other indices of malingering.

**Dissimulation on Diagnostic Screening Instruments**

Screening for psychiatric impairment is a concern for many institutions including jails, prisons, emergency rooms, and crisis centers. As mental health resources continue to decrease and regulations regarding the administration of services become increasingly bureaucratized (e.g., Health Management Organizations as discussed by Durham, 1994), the rapid and accurate assessment of psychiatric symptomatology has become a priority. However, the accuracy of mental health screening measures is limited by the lack of research on their fakability. In light of the brevity and high face validity of many of the existing screening measures (e.g., the RDS, SPS, Beck Depression Inventory) the likelihood of undetected feigning is substantial. The cost of this oversight is high in that either a comprehensive, time-intensive evaluation will be conducted or inappropriate/unnecessary treatment may be administered.

The results of the current study suggest that the RDS, SPS, and the SADS-C are vulnerable to faking. The finding that participants were able to alter successfully their
presentations is cause enough for concern. However, when considered in light of the current trends in mental health care (e.g., reduction of resources and funding) the potential impact of these results is particularly distressing. Specifically, if mental health professionals base decisions primarily on screening measures, it is very likely that malingerers will be successful at gaining access to limited mental health resources. In the current sample, 95.7% of the suspected malingerers would have received treatment on the basis of screening instruments alone.

A conflict exists between the demands associated with adequate provision of treatment and the limits imposed by the existing mental health care system. The development of the RDS was one of the first attempts to address the need for psychiatric assessment that is both cost-effective and efficient (Teplin & Schwartz, 1989). The results of this study and others (see, Hart et al., 1993; Rogers et al., 1995) suggest that the RDS is an effective screen for severe psychopathology but its discriminant validity is very limited. In relation to correctional settings, the RDS is constrained by (a) an extremely limited range of symptoms, (b) the lack of evaluative criteria for highly problematic behaviors (e.g., prior suicide attempts or current suicidal ideation) and (c) its face validity.

As already described, the RDS was designed as a first-stage screen for severe mental illness in a large urban jail
facility. The current application of the RDS addressed its effectiveness with screened individuals. Thus, the RDS functioned as a second-stage screen in a sample that is less heterogenous with respect to severe mental illness. The change in composition of the sample provided in a more stringent test of its discriminatory power since the base rate for severe mental illness was likely to be substantially higher than that of the original validation study.

The utility of the RDS as a second-stage screen was not supported by the current study. The application of the most liberal cut scores resulted in 78% (n = 26) of the sample being referred for further psychiatric evaluation. This included 95.8% of the suspected malingerers, 40.3% of the patients, and 100% of the indeterminant group. Inspection of the SPS and the SADS-C scores indicated that many of the individuals screened out by the RDS may have been currently experiencing symptomatology that required treatment. Specifically, six individuals scored in the clinical range on the SADS-C (i.e., symptoms of depression, mania, or psychosis) and 16 met the SPS criteria for further evaluation (T-score > 60 or < 40). In summary, the RDS's limited range of symptoms hinders its ability to detect psychiatric impairment.
Screening for Suicidality

Evaluation of suicide risk is a critical component of mental health evaluation in jail facilities. It is important not only because the base rate of mental illness in jail facilities is high but because the physical conditions and the psychosocial stressors related to incarceration often exacerbate already existing problems.

Comprehensive evaluation of suicide risk is a very difficult, time-consuming task that requires knowledge of the patient's past and present psychiatric functioning. As mentioned previously, the nature of the screening process prohibits a time-intensive approach and requires that the clinician adopt a different strategy to obtain accurate information. Specifically, each aspect of the evaluation must be conducted in an effective and cost-efficient manner with each instrument adding new information. One method of accomplishing this goal is to use a brief battery of screening instruments with each contributing to the incremental validity of the evaluation process.

In the current study, the SPS was the only measure developed specifically to evaluate suicide probability. For clinical purposes, this measure was used in conjunction with the SADS-C to assist in the evaluation of current symptoms, their severity, and associated risks for self-harm. Individuals who were considered to be at high risk for self-harm were immediately brought to the attention of the staff.
The remaining evaluations were completed within three days thereby providing TXMHMR staff with an initial evaluation of the inmates current psychiatric status.

Like other screening measures, the SPS was presumed to be easily falsifiable but this supposition had not been empirically tested. Thus, it is important to know how this response style impacts the results of the measure and be able to relate these results to treatment considerations in correctional settings. Since the risks associated with obtaining false-negative (missed suicidal intentions) findings are too great to warrant decisions based on screening measures alone (i.e., identifying an individual as a malingering when they are truly suicidal), the information gleaned from an initial screening battery would likely be most important for future evaluations. For example, the information may be relevant for detainees who undergo further evaluation of competency to stand trial or criminal responsibility or who request an evaluation for medication or special services while in the jail setting.

As previously described, to interpret the SPS a clinician must decide on level of presumptive risk before treatment recommendations can be established. The treatment recommendations vary on a number of dimensions including extent of services, intensity of services, and restriction of civil liberties. In general, higher levels of presumptive risk require more extreme interventions.
In the current study, level of presumptive risk had little effect on the outcome for suspected malingerers. Regardless of whether a high or intermediate level of presumptive risk was applied, suspected malingerers were classified as being at severe or moderate suicide risk. According to the SPS manual, severe suicide risk requires that "extreme suicide precautions" be implemented including immediate hospitalization, one-on-one monitoring at all times, and use of psychotropic medications. Alternately, the patient group was classified as either a mild or subclinical suicide risk depending on the level of presumptive risk chosen. Suggestions for treatment at these levels include further evaluation, therapeutic intervention, and routine monitoring. It is apparent that had the guidelines of the SPS been followed, many of the suspected malingerers would have placed on significant suicide precautions.

Perhaps the most disconcerting aspect of these findings is that suspected malingerers not only chose to feign suicidality but that they appear to be adept at this task. In fact, the results of the current study closely match those found for the cross-validation sample at the test construction stage (Zachary, Roid, Cull, & Gill, 1982 as cited in Cull & Gill, 1988). The suspected malingerers produced mean scores that were very similar to those obtained by patients who had attempted suicide a maximum of
two days before completing the SPS. The patient group, on the other hand, produced results that were similar to the psychiatric inpatient sample who had no history of suicidal behavior. The ease with which suicidality can be feigned is of great concern since malingerers and bona fide patients are likely to be in competition for very limited mental health resources in jail facilities.

**Assessment of Malingering**

The fundamental assumption of this and other studies of malingering, is that obtaining honest and accurate information is integral to the assessment of mental illness. Veracity of self-report assumes paramount importance when decisions regarding service delivery and protection of life rely heavily on verbal statements of distress and impairment. This section presents arguments for and against the use of various strategies in the assessment of malingering. The discussion begins with an evaluation of the use of inconsistent responding as a hallmark of malingered response style. Following from this, the utility of the SADS models of malingering and the SADS scales are presented in terms of validation and clinical utility. The section concludes with a discussion of the clinical utility of the M Test.

**Inconsistent Responding**

Inconsistent responding is one of the most commonly discussed signs of a malingered presentation (see, Cornell &
Hawk, 1989; Ogloff, 1990; Rogers, 1988; Ziskin, 1984). Typically response consistency has been discussed in relation to symptom profiles. Specifically, malingered presentations are thought to be characterized by symptom profiles that (a) do not match known diagnostic entities, (b) include contradictory symptom pairs, and (c) are inconsistent with collateral information (e.g., behavioral observation, interviews with family) (see, Cornell & Hawk, 1989; Davidson, 1949; Ossipov, 1944; Rogers, 1984).

As described by Rogers (1988), the SADS models of malingering address all three of the aforementioned characteristics related to an inconsistent response set. Specifically, Unusual Symptom Combinations provides a method of evaluating "unknown diagnostic entities", Contradictory Symptom Pairs provides a direct measure of "contradictory symptoms" and Part B of the Clinical Decision Model, not addressed this study, evaluates malingering via ancillary sources or admissions of malingering.

The results of the discriminant function analyses of the SADS and the SADS-C models suggest that inconsistency is an important predictor of suspected malingering in the current sample. In both analyses, Contradictory Symptom Pairs and Unusual Symptom Combinations added to the predictive accuracy of the models. The hit rate for both models was greater than 90% with the Rare Symptom Scale, Contradictory Symptom Pairs, Unusual Symptom Combinations,
and Symptom Severity Scale (number of items scored ≥ 3 or ≥ 5) included in both functions. Each function accounted for slightly more than 50% of the variance.

A second method of examining consistency is to evaluate an individual's ability to respond similarly across two or more time periods. This methodology is different from inconsistent symptom profiles in that the question being asked is "Did the individual respond in a reliable manner?" rather than "Does this symptom profile make clinical sense?" According to Greene (1988), a consistent response set "indicates only that the patient has responded consistently but not necessarily accurately" (p. 132). Following from this statement, it is easy to see that an inconsistent response set would indicate only that the individual responded in an unreliable manner. Thus, reliability, in and of itself, provides limited information with respect to malingering. However, lack of reliability can be used as corroborative evidence to support a finding of malingering that is based on multiple sources of information. For example, malingerers may produce contradictory information over time or become too tired or confused to maintain a consistent response set (Resnick, 1993; Ziskin, 1984).

The SADS and the SADS-C provided an excellent opportunity to address whether consistency of item endorsement is an effective predictor of malingered responding. Unlike multiscale inventories, where evaluation
of consistency occurs on one instrument, the design of the study allowed for re-evaluation of the same items following a brief time interval. The fact that the items of the SADS-C are imbedded within the SADS provided protection against consistency related to rote memory since the SADS-C items are dispersed throughout the SADS Part 1 and not presented in the same sequence.

On a statistical level, the results of the current study support the hypothesis that, on average, suspected malingerers respond less consistently over time than do patients. On a clinical level, however, this variable did not appear to be helpful in the identification of suspected malingerers since the degree of inconsistency was too small to be of value. To illustrate, the average difference score for malingerers on the 27 items did not exceed 1.82 (SD = 1.68; see, Table 7) and total scores on the sub-scales of the SADS-C were within two points 80% of the time (see, Table 8).

**Concurrent Validity of SADS Models of Malingering**

_SADS Models of Malingering._ The empirically derived SADS models of malingering (Rogers, 1988) performed in accordance with their respective goals. The Threshold Model performed well as a screening measure retaining 100% of the suspected malingerers and 42% of the patients for further evaluation. This model was designed for application in settings where sensitivity is of primary concern and the
SADS based assessment of malingering is used in conjunction with other measures of malingering.

Like the Threshold Model, the Clinical Decision Model performed in accordance with the design specification. The model was intended to be very conservative in its evaluation of malingering with the primary concern being positive predictive power rather than sensitivity. In the current study, 54.5% of the suspected malingerers were misclassified as patients and PPP was .83. However, it appears that the model may be too conservative in the assessment of malingering. The goal would be to identify new cut scores that would provide a more balanced classification where the trade off between sensitivity and specificity is not so great and positive predictive power is not unduly compromised.

In fact, the results of the discriminant function provide support for the notion that more effective cut scores are possible. When compared to the Clinical Decision Model of the SADS, the discriminant function analysis improved classification with respect to all parameters. The values represented a "middle ground" between the Threshold and Clinical Decision Models where the choice of cut scores is directly tied to perceived consequences associated with an error in judgement (see Table 21).
Table 21

Classification Accuracy of SADS Models of Malingering

<table>
<thead>
<tr>
<th>Models</th>
<th>Threshold</th>
<th>Clinical Decision</th>
<th>Discriminant Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hit rate</td>
<td>70.8%</td>
<td>81.9%</td>
<td>91.7%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>1.0</td>
<td>0.46</td>
<td>0.77</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.58</td>
<td>0.96</td>
<td>0.98</td>
</tr>
<tr>
<td>PPP</td>
<td>0.51</td>
<td>0.83</td>
<td>0.94</td>
</tr>
<tr>
<td>NPP</td>
<td>1.0</td>
<td>0.80</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Of course, the use of step-wise discriminant function analysis is somewhat controversial since "the order of entry may be dependent on trivial differences in relationships among predictors in the sample that do not reflect population differences" (Tabachnick & Fidell, 1989, p. 531). The results of the current study need to be replicated to ensure a stable function that represents population differences.

The finding that rare symptoms are not reliable predictors of malingering (Rogers, 1997) was not supported in the current study. Overendorsement of rare symptoms has been considered one of the classic signs of malingering and has been supported numerous times in the literature.
(Dahlstrom, Welsh, & Dahlstrom, 1972; Resnick, 1993). The results of the current study provide further support for the utility of rare symptoms in the evaluation of malingering. Not only were the rare symptoms scales on the SADS and the SIRS moderately correlated (suggesting that the two scales evaluate similar constructs) but the classification accuracy of the SADS models of malingering decreased when the Rare Symptom Scale was removed. It is speculated that the lack of discriminatory power noted by Rogers (1997) is the result of an inappropriate comparison sample rather than a real effect. Rogers compared SADS data across three samples: Two forensic samples and one non-forensic, out-patient psychiatric sample. It is hypothesized that the out-patient sample may have been disengaged from the interview process due to high level of symptomatology that interfered with their ability to attend. The nature of the sample differences cannot be directly evaluated since the out-patient sample was not screened for response style.

**SADS-C Models of Malingering.** The current study is the first to evaluate the generalizability of the SADS models of malingering to the SADS-C. The results suggest that the models are generalizable to the SADS-C and that the Threshold Model performs slightly better on the SADS-C than the SADS. The applicability of the models to the SADS-C provides preliminary support for the validity of the models. However, there is some indication that the models may
function differently on each of the measures and that these differences may be related to scale properties rather than questionable construct validity.

As already discussed, the Rare Symptoms scale was significantly altered in response to content of the SADS-C; the SADS-C Rare Symptoms scale includes only three of the 10 original items. Not surprisingly, the Rare Symptoms scale was not found to be an effective predictor in the classification of suspected malingering (i.e., the classification statistics remained unchanged when Rare Symptoms was removed from the function). However, the removal of the Rare Symptoms scale revealed an unexpected finding that is important to the understanding of the SADS-C models of malingering. Specifically, the removal of this scale coincided with the statistical exclusion of Contradictory Symptom Pairs from the function.

As previously described, Contradictory Symptom Pairs provides a measure of the degree to which a person presents a symptom picture that makes clinical sense. The statistical exclusion of this scale suggests that suspected malingers and patients do not differ with respect to this variable on the SADS-C. In fact, going back to the discussion of inconsistency in general, this finding supports the notion that consistency provides little independent information with regard to response style. It is hypothesized that this finding relates only to the SADS-C
and that Contradictory Symptom Pairs is an effective predictor of malingering in a SADS based assessment. The rationale is that the shortened interview and brief period of evaluation make it easier for both patients and suspected malingerers to present in a consistent or reliable manner. For example, the SADS-C questions regarding lack of energy and increased energy are only separated by nine questions on the SADS-C. Thus, it is hypothesized that the structure of the SADS-C prevents this response style from being an effective strategy. Interestingly, all of the classification statistics remained the same despite deletion of two of the four variables. It appears that, on the SADS-C, Unusual Symptom Endorsement and Over-endorsement of Symptoms were sufficient for predicting malingering.

Malingering Scales on the SADS and SADS-C

The development of malingering indices on the SADS and SADS-C was an exploratory procedure relying solely on empirical findings. The items included in each scale were chosen exclusively on their ability to predict suspected malingering. Each scale was scored in two ways: (a) summation of severity ratings and (b) dichotomous scoring based on presence or absence of symptoms. At the conceptual level, the summing method is believed to reflect overendorsement of symptoms and symptom severity. The more items endorsed and the more severe the rating, the higher the total score. Dichotomous scoring, on the other hand, is
believed to function in a manner similar to previously
developed rare symptom scales. Like the rare symptoms
scales, few patients endorsed more than one or two of the
items. Tests of reliability and homogeneity provide initial
support for the psychometric stability of the indices.

Overall, the two scales demonstrated excellent
classificatory accuracy when differentiating between
suspected malingerers and patients. Hit rates were
approximately 90% with sensitivity and specificity greater
than .80. Optimal cut scores were easily identified using
both scoring procedures. Only the SADS-C-Sum, the scale
which takes into account overendorsement and symptom
severity, produced two equally appropriate cut scores which
differ with respect to classification rates. If the
clinician/researcher desires to reduce the rate of false
positives, a stringent cut score that maximizes hit rate,
specificity, and positive predictive power is provided. On
the other hand, if the clinician/researcher is interested in
reducing false negatives, a more liberal cut score that
maximizes hit rate, sensitivity, and negative predictive
power is recommended.

The development of screening procedures for malingering
on the SADS and SADS-C is beneficial in two ways. First, as
already discussed, structured interviews provide a high
level of control over examiner variability. Thus, when
using either of these instruments, there is less concern
regarding the potential confounds associated with examiner style or preconceptions. Second, both the SADS-C and the SADS have been found to be effective in the evaluation of psychopathology. Thus, screening for malingering can be conducted while simultaneously obtaining necessary diagnostic information. If malingering is not a concern, then the evaluation is beneficial in terms of the psychiatric information obtained. If, on the other hand, the scales suggest that feigning is a possibility, the clinician/researcher benefits from the knowledge that a more comprehensive evaluation of malingering may be warranted. Clearly, the scales are not recommended as definitive tests of malingering. They are designed strictly as screening indices and should always be used in conjunction with other measures of malingering including the Threshold and Decision models of the SADS.

At the present time, the scales are recommended for use as screens to be used on conjunction with multiple indices of malingering. Both scales require extensive evaluation regarding their reliability and validity in different settings, with different populations, using different methodology. The need for further validation is of paramount importance since these scales are developed using strictly empirical procedures (Briggs & Cheek, 1989; DeVillis, 1991). Such scales are likely to be biased in favor of the research sample as the strategy for item
selection maximized sensitivity and specificity. Furthermore, the use of discriminant function analysis in the final stage of item selection further limits the generalizability of the scales. Specifically, the methodology inherent in a stepwise discriminant function analysis eliminates viable items on the basis of shared variance (Tabachnick & Fidell, 1989). Consequently, if the procedure was repeated using a new sample, it is possible that item composition of the scales would be different. Despite the inherent limitations of empirical scale development, the scales appear to be psychometrically sound and show promising predictive validity such that further investigation is warranted.

The M Test: A Viable Screening Measure

The results of the current study suggest that the modified version of the M Test is a viable screening measure for malingering in a forensic population. The classification rates closely match those of the original validation study (Rogers, Bagby, & Gillis, 1992) and are better than those found by Smith, Borum, and Schinka (1993). In contrast to the Smith et al study, the error rate for both Option A (16.3%) and Option B (23.3%) were lower than the base rate of malingering in the current sample (27.3%). This finding indicates that the M Test predicts malingering at a rate better than chance. As previously mentioned, the results of Hankins and colleagues (1993) study provided only
minimal support for the validity of the Rule-In and Rule-Out scales and raised four primary concerns in relation to test validity. These concerns were addressed in the current study and discussed in the following paragraphs.

First, unlike the original validation study, the gold standard for malingering was a general measure of malingering (i.e., the SIRS) rather than one or more malingered psychotic symptoms. One of Hankins et al.'s (1993) criticisms of the Rogers et al. study was that external validity may have been compromised due to biased sampling that favored malingered psychosis. In the current study, this problem was controlled for by the use of a gold standard that is not diagnosis-specific.

The second criticism rendered against the Rogers study related to a supposed 50% base rate of malingering. The base rate of suspected malingering in the current sample was either 19.7% (total sample) or 27.3% depending on the groups used in the analysis. In fact, base rate of malingering in the current study was very similar to that of the Rogers et al. study which was 23.6% (not 50% as reported by Hankins et al.). Thus, neither the current study nor that of Rogers et al., maximized the predictive power of the M Test by evaluating its efficacy on a sample with an atypical base rate of suspected malingering.

Third, Hankins et al. stated that the use of "extreme" malingerers would exaggerate group differences resulting in
a strong bias in favor of the validity of the M Test. The authors suggested that, since malingering is not conceptualized as a dichotomous variable, comparisons should be conducted with a sample that represents the "continuum of malingering". This methodology was incorporated in the current study and the M Test was shown to be effective as a screening device and appeared to classify in a manner similar to the SIRS. Specifically, using the less stringent cut score (Option A), 48.5% of the indeterminant group was retained for further evaluation of malingering. The more stringent cut score (Option B), which is recommended for use in forensic settings, retained 69.7% of the indeterminant group for further evaluation of malingering. On average, individuals who were retained for further evaluation produced higher profiles on the SIRS than did those who were eliminated.

Lastly, Hankins and colleagues stated that the findings of the Rogers et al. study must be cross-validated before the scoring procedure can be recommended for clinical use. The current study serves as a cross-validation study for the Rule-in/Rule-out method and supports the validity of the scoring procedure. Classification of patients and suspected malingerers was virtually identical in both studies providing strong support for the validity of the method in forensic populations.
Developing a Screening Battery for Malingering and Mental Illness

The development of efficient screening procedures is a very complex and difficult task. As discussed by Teplin and Schwartz (1989) the relative costs of making Type I and Type II errors must be considered at each step of the process. In terms of screening procedures, the costs of a Type I error (referring an individual for further evaluation when they are not truly ill) include an increase in the need for services (e.g., staff and resources), inconvenience to the examinee, and possible neglect of a bona fide patient who presents as less seriously ill. The cost of a Type II error (not referring an individual for further evaluation or treatment when they are truly ill) relates primarily to negligence with regard to provision of services. For example, a bona fide patient who remains untreated may engage in self-destructive behavior, act out violently, and/or disrupt normal institutional functioning. Thus, it appears that a two-stage evaluation process, beginning with administration of the RDS and the SPS, may be a cost-effective and reliable way of screening for mental illness. At the second stage of evaluation, administration of a measure of current symptomatology and malingering are considered necessary components.

Initial validation of the two-stage screening procedure indicated that this screen may provide jail personnel with a
more comprehensive procedure that is both cost and time efficient. Stage 1 of the screen involves administration and scoring of the RDS and the SPS. The time required for such an evaluation is between 15 to 20 minutes. However, the required time can be substantially reduced with group administration of the SPS and the utilization of a computerized scoring procedure. In the current study, 10.7% (8 patients and 1 suspected malingerer) of the sample was eliminated from further evaluation. If this procedure were applied to the current jail setting (i.e., TXMHMR-TCJ) who receive 800 referrals per month, approximately 86 persons would be eliminated from further evaluations from second-stage evaluations. Exact calculation of the number of individuals screened out at the first stage of evaluation is not possible but is expected to be substantial. An additional benefit to thorough screening at intake is the reduction of false negatives (i.e., misidentification of patients as non-patients). It was noted that in the current study, a large number of inmates (both participants and non-participants) stated that they were not evaluated properly at intake and that this resulted in a delay of treatment or lack of treatment/evaluation all together.

Stage 2 of the assessment process involved the administration of the SADS-C and the M Test respectively. It is estimated that, on average, the time required to administer and score both instruments would be 25 minutes.
Once again, the time involvement could be significantly reduced with group administration of the M Test. Since this is a second-stage screening process, it is likely that mental health professionals would be conducting the evaluation and would require a moderate amount of training on the administration of the SADS-C. Interpretation begins with the sub-scales of the SADS-C and continues with an evaluation of malingering using the SADS-C-Sum. Individuals who meet criteria for mental illness and are suspected to be malingering are then evaluated with the M Test Rule-In/Rule-Out procedure.

Interpretation of the SADS-C places an individual in one of the following three categories: Suspected malingerer, patient, or non-patient category. In the current sample 32.1% of the sample were classified as non-patient, 23.8% as patient, and 33.4% as possible malingerers. Application of the M Test Rule-In/Rule-Out procedure screened out 5 of the 7 remaining patients and retained all of the remaining malingerers (n = 21).

The procedure has been delineated as a two-stage process on the basis of test administration and a three stage process on the basis of interpretation. However, there is no reason why the M Test could not be administered at intake along with the RDS and SPS. In fact, this may be beneficial in regards to time since the M Test and the SPS are paper and pencil measures completed by the detainee.
Interpretation of the measures would follow the same guidelines with the results of the M Test passed forward to the second-stage evaluator.

In conclusion, the recommended screening procedure is believed to be a cost effective way to adequately screen for mental illness in a jail facility. The additional time commitment on the part of jail personnel is not substantial and comes with the advantage of a psychometric screen for suicide risk. As previously discussed, addressing suicide risk at booking is imperative since the majority of suicides occur in the first 24 hours (Rakis, 1984; Rood & Faison, 1988).

Limitations to the Current Study

1. The Clinical Decision Model of the SADS was not tested exactly as delineated by Rogers (1988). Specifically, none of the individuals were asked if they had maligned their responses and clinical evidence based on observation or other psychometric data were not used in conjunction with the SADS to make a final decision. These two criteria are important components to the model but could not be evaluated in the current study. It is not known what impact this information would have had on the classification rates in the current study.

2. The use of a pre-selected sample of jail referrals prevents a complete understanding of the effectiveness of the recommended screening procedure. In a jail setting, the
RDS and the SPS are likely to be most beneficial at the time of booking. The information gleaned from the current study was that only 10.7% of the recommendations were potentially unnecessary. However, information is not available in the current study on how many patients were missed by the current referral procedures.

3. The newly developed malingering scales on the SADS and the SADS-C require further validation regarding reliability and validity. In addition to cross-validating the current scales, it is particularly important to test the effects of other response patterns (i.e., random responding and haphazard responding) as well as attempts to simulate specific disorders (e.g., bipolar disorders, major depressive disorder). Empirical scale development may render these indices particularly vulnerable to response patterns.

4. The classification data from the SADS models of malingering are most germane when using data obtained directly from patients. In the current study, the effects of other data sources on the SADS responses was not evaluated. Therefore, the standard use of the SADS integrating ancillary sources (e.g., family member), may alter some of the responses, increasing or decreasing the reported symptomatology. If malingering is suspected, clinicians may wish to make ratings solely on the patient's
self-report and modify these ratings once the patient is screened for possible malingering.
NOTES

1The State of New York, Commission of Corrections, Office of Mental Health developed the Suicide Prevention Screening Guidelines (SPSG), a standardized form that was implemented in correctional facilities statewide. The form is designed to be administered within the first 72 hours of incarceration and requires approximately 3 to 5 minutes to administer (see Sherman & Morschauser, 1989, for more details and a sample copy).

2The PAI was administered for purposes external to the current study.

3The first stage of the evaluation was conducted by one of two advanced graduate students in clinical psychology (Randy Salekin, M. S. and Vianey Reinhardt, M. S.). The second stage of the evaluation was conducted by the primary investigator (Karen Ustad, M. S.).

4The discriminating variables used were rare symptoms, contradictory symptom pairs, unusual symptom combinations, and symptom severity (≥ 5). Overendorsement of symptoms could not be included in these analyses due to an unacceptably high correlation with symptom severity (.80). Refer to Appendix E for the discriminant analysis using overendorsement of symptoms in combination with the three other variables.
The discriminant function for the SADS model is as follows: score = -0.6568434 + (rare symptoms)(0.8226612) + (contradictory symptom pairs)(0.4798181) + (unusual symptom combinations)(0.3609589) + (number of symptoms scored ≥ 5)(-0.0697769). This value should be compared to the mid-point between group centroids which was 0.426285. Scores equal to or above this value are considered to be statistically similar to the malingering group and those below are statistically more similar to the patient group.

The discriminant function for the SADS models when rare symptoms were not included as a predictor variable is as follows: score = -.7585677 + (contradictory symptom combinations)(.7340416) + (unusual symptom combinations)(.3982510) + (endorsement of symptoms ≥ 5)(-.03699452). This value should be compared to the mid point between the group centroids which was .341955. Scores equal to or above this value are considered to be statistically similar to the suspected malingerers and those below are statistically similar to the patient group.

The predictor variables for the SADS-C models were rare symptoms, unusual symptom combinations, contradictory symptom pairs, and overendorsement of symptoms (≥ 3). The analyses were repeated using severity of symptomatology instead of overendorsement of symptoms. These results are reported in Appendix F.
The discriminant function for the SADS-C model is as follows:  
\[ \text{score} = -1.538226 + (\text{count of rare symptoms})(-0.8931123) + (\text{count of contradictory symptom pairs})(0.3495140) + (\text{count of unusual symptom combinations})(0.08747134) + (\text{count symptoms scores} \geq 3)(0.2036390). \] 
This value should be compared to the mid point between the group centroids which was .566705. Scores equal to or above this value are considered to be statistically similar to the suspected malingering group and those below are statistically similar to the patient group.

The discriminant function for the SADS-C models when rare symptoms were not included as a predictor variable is as follows:  
\[ \text{score} = -1.564297 + (\text{count of unusual symptom combinations})(0.08448774) + (\text{count of the number of symptoms scored} \geq 3)(0.2010962). \] 
This value should be compared to the mid point between the group centroids which was 0.54911. Scores equal to or above this value are considered to be statistically similar to the suspected malingerers and those below are statistically similar to the patient group.

Clinical range on the SADS-C was defined as follows: 
\[ \geq 34 \text{ on the depression scale,} \geq 10 \text{ on the mania scale, and} \geq 9 \text{ on the paranoia scale.} \] 
Cut scores were developed from research conducted by Johnson, Magaro and Stern (1986). The cut scores = the mean score obtained by each patient group minus one standard deviation.
The final results of the screen remained identical when using SADS-C-Count cut score of $\leq 1$ at stage 2 (retain if score is two or greater). The classification rates at stage 2 change slightly when applying this cut score. The number of individuals retained for further evaluation increases from 7 to 11 while classification of suspected malingerers remained the same.
APPENDIX A

SYMPTOM PATTERNS FOR THE SADS DECISION MODELS
Symptom Patterns for the SADS Decision Models

1. Items for the Rare Symptoms Scale
   
   #354 Elevated or expansive mood (past week)
   #356 Needs less sleep than usual (past week)
   #360 Increase in goal directed activity (past week)
   #427 Thought withdrawal
   #430 Delusions of guilt or sin
   #432 Somatic delusions
   #475 Loosening of associations
   #476 Impaired understandability (past week)
   #514 Poverty of content of speech
   #515 Neologisms

2. Contradictory Symptoms
   
   #234-#353 Subjective feelings of depression
   Elevated or expansive mood
   
   #242-#361 Negative evaluation of self
   Grandiosity
   
   #272-#314 Insomnia
   Sleeping more than usual
   
   #315-#357 Subjective feeling of lack of energy
   Unusually energetic
   
   #317-#320 Poor appetite
   Increased appetite
   
   #326-#359 Pervasive loss of interest
   Increase in goal directed activity
   
   #334-#342 Agitation
   Psychomotor retardation

Appendix A continues
Appendix A continued

3. Unusual Symptoms Combinations

- #317-#239  Poor appetite & Worrying (past week)
- #317-#243  Poor appetite & Negative evaluation of self (past week)
- #317-#245  Poor appetite & Discouragement (past week)
- #317-#266  Poor appetite & Psychic anxiety (past week)
- #317-#331  Poor appetite & Subjective anger (past week)
- #317-#419  Poor appetite & Suspiciousness (past week)
- #334-#245  Agitation & Discouragement (past week)
- #334-#266  Agitation & Psychic anxiety (past week)
- #428-#239  Persecutory delusions & Worrying (past week)
- #428-#245  Persecutory delusions & Discouragement (past week)
- #428-#313  Persecutory delusions & Trouble sleeping (past week)
- #428-#331  Persecutory delusions & Subjective anger (past week)
APPENDIX B

MODIFIED SYMPTOM PATTERNS FOR THE SADS-C DECISION MODELS
Modified Symptom Patterns for the SADS-C Decision Models

1. Rare Symptoms

Item 235 Elevated or expansive mood (past week)
236 Needs less sleep than usual (past week)
238 Increase in goal directed activity (past week)

2. Contradictory Symptoms

Items 213-235 Subjective feelings of depression
Elevated mood or optimistic attitude
Items 216-239 Negative evaluation of self
Grandiosity
Items 227-237 Subjective feeling of lack of energy
Unusually energetic
Items 230-238 Pervasive loss of interest
Increase in goal directed activity
Items 233-234 Agitation
Psychomotor retardation
Items 217-235* Discouragement
Elevated mood or optimistic attitude
Items 223-236* Insomnia
Need less sleep than usual

Item #314 and #320 on the SADS are not present on the SADS-C therefore two contradictory symptom pairs could not be included. The asterisk denotes new symptom pairs.

3. Unusual Symptoms Combinations

Items 228-214 Poor appetite & Worrying

Appendix B continues
Appendix B continued

<table>
<thead>
<tr>
<th>Items</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>228-216</td>
<td>Poor appetite &amp; Negative evaluation of self</td>
</tr>
<tr>
<td>228-217</td>
<td>Poor appetite &amp; Discouragement</td>
</tr>
<tr>
<td>228-220</td>
<td>Poor appetite &amp; Psychic anxiety</td>
</tr>
<tr>
<td>228-231</td>
<td>Poor appetite &amp; Subjective anger</td>
</tr>
<tr>
<td>228-240</td>
<td>Poor appetite &amp; Suspiciousness</td>
</tr>
<tr>
<td>233-217</td>
<td>Agitation &amp; Discouragement</td>
</tr>
<tr>
<td>233-220</td>
<td>Agitation &amp; Psychic anxiety</td>
</tr>
<tr>
<td>241-214</td>
<td>Delusions &amp; Worrying</td>
</tr>
<tr>
<td>241-217</td>
<td>Delusions &amp; Discouragement</td>
</tr>
<tr>
<td>241-223</td>
<td>Delusions &amp; Insomnia</td>
</tr>
<tr>
<td>241-231</td>
<td>Delusions &amp; Subjective anger</td>
</tr>
</tbody>
</table>

**Note.** #241 on the SADS-C is not identical to #428 on the SADS but is very similar and was used for the purposes of this scale.
APPENDIX C

SADS SCALE FOR SUSPECTED MALINGERING
### SADS Scale for Suspected Malingering

<table>
<thead>
<tr>
<th>Item Number</th>
<th>Symptom</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. #442</td>
<td>Auditory hallucinations</td>
</tr>
<tr>
<td>2. #365</td>
<td>Motor hyperactivity</td>
</tr>
<tr>
<td>3. #348</td>
<td>Psychomotor retardation (past week)</td>
</tr>
<tr>
<td>4. #443</td>
<td>Voices commenting on behavior</td>
</tr>
<tr>
<td>5. #341</td>
<td>Agitation (past week)</td>
</tr>
<tr>
<td>6. #450</td>
<td>Severity of hallucinations</td>
</tr>
<tr>
<td>7. #329</td>
<td>Depersonalization</td>
</tr>
<tr>
<td>8. #239</td>
<td>Worrying (past week)</td>
</tr>
<tr>
<td>9. #455</td>
<td>Sensorium (past week)</td>
</tr>
</tbody>
</table>

**Note:** The items in the scale are listed in the order that they entered the discriminant function. At step two, item #235 (depression rated for the past week) was entered into the function and was removed at step seven.
APPENDIX D

SADS-C SCALE FOR SUSPECTED MALINGERING
### SADS-C Scale for Suspected Malingering

<table>
<thead>
<tr>
<th>Item Number</th>
<th>Symptom</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>#230 Loss of interest</td>
</tr>
<tr>
<td>2.</td>
<td>#240 Trouble with people</td>
</tr>
<tr>
<td>3.</td>
<td>#220 Psychic anxiety</td>
</tr>
<tr>
<td>4.</td>
<td>#222 Obsessions/compulsions</td>
</tr>
<tr>
<td>5.</td>
<td>#217 Discouragement</td>
</tr>
<tr>
<td>6.</td>
<td>#242 Presence of hallucinations</td>
</tr>
</tbody>
</table>

*Note: The items in the scale are listed in the order that they entered the discriminant function.*
APPENDIX E

DISCRIMINANT FUNCTIONS FOR GROUP CLASSIFICATION

FOR THE SADS USING CLINICAL ENDORSEMENT
Discriminant functions for group classification for the SADS using clinical endorsement (count of the number of symptoms endorsed > 3) instead of severity of symptoms (count of the number of symptoms endorsed > 5).

<table>
<thead>
<tr>
<th>Predicted Group</th>
<th>Suspected Malingerer</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspected Malingerer</td>
<td>18 (85.7%)</td>
<td>3 (14.3%)</td>
</tr>
<tr>
<td>Patient</td>
<td>4 (7.8%)</td>
<td>47 (92.2%)</td>
</tr>
</tbody>
</table>

Wilks Lambda = .46

χ² (2, η = 72) = 52.96, p < .001

Canonical Correlation = .73

Hit Rate = 90.3%

Sensitivity = .82

Specificity = .94

Positive Predictive Power = .86

Negative Predictive Power = .93
APPENDIX F

DISCRIMANT FUNCTIONS FOR GROUP CLASSIFICATION
FOR THE SADS-C USING SEVERITY OF SYMPTOMS
Discriminant functions for group classification for the SADS-C using severity of symptoms (count of the number of symptoms endorsed ≥ 5) instead of clinical endorsement (count of the number of symptoms endorsed ≥ 3).

<table>
<thead>
<tr>
<th>Actual Group</th>
<th>Suspected Malingerer</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspected</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>Malingerer</td>
<td>(77.7%)</td>
<td>(33.3%)</td>
</tr>
<tr>
<td>Patient</td>
<td>10</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>(14.3%)</td>
<td>(85.7%)</td>
</tr>
</tbody>
</table>

Wilks' lambda = .56

χ² (2, n = 88) = 49.82, p < .001

Canonical correlation = .67

Hit Rate = 84.1%

Sensitivity = .58

Specificity = .94

Positive Predictive Power = .78

Negative Predictive Power = .86
REFERENCES


Bowring v. Godwin, 551 F.2d 44 (4th Cir. 1977).


need for common approaches and international perspectives.  


