THE EFFECT OF SYMPTOM INFORMATION AND VALIDITY
SCALE INFORMATION ON THE MALINGERING OF
DEPRESSION ON THE MMPI-2

by

GINA L. WALTERS, B.A., B.F.A., M.A.

A DISSERTATION

IN

PSYCHOLOGY

Submitted to the Graduate Faculty
of Texas Tech University in
Partial Fulfillment of
the Requirements for
the Degree of

DOCTOR OF PHILOSOPHY

Approved

Accepted

August, 1998

ACKNOWLEDGMENTS

Completing my dissertation would not have been possible without the encouragement and support of many people. First and foremost, I would like to thank my dissertation chair, Jim Clopton, for devoting so much time and energy (not to mention his office) to this project. But mostly, I would like to thank Dr. Clopton for his unwavering support, encouragement, and guidance during my five years at Texas Tech. He is truly a fantastic mentor and friend.

There are several people who went beyond the call of duty in helping me to obtain a clinical sample for this study. I want to thank Jim Butcher and John Graham for graciously agreeing to share their data set with me. I feel extremely lucky to have had access to this data set. Additionally, I am greatly appreciative of the help of Steve Finn and Pitts Davis in my data collection efforts.

I would like to extend a special thanks to Dennis Edwards at the Texas Tech ATLC who went out of his way to help me through several computer "crises." His creation of a custom-made computer program (which was quite time-consuming for him) was an incredible time saver, allowing me to meet several crucial deadlines.

I would also like to thank all the people who helped me collect and score data (Katie Norton, Cathy Thompson, Kenda Morgan, and Veronika Polisensma).

Last, but not least, I would like to thank my all my friends and family for their support during the dissertation process. I want to thank my loyal friend, Katherine Green, who consistently offered encouragement and support during the good times and bad. I want to express my gratitude to my sister, a seasoned dissertation veteran, who provided much guidance and reassurance. I am most indebted to my parents who have "rooted me on" throughout my entire educational process (yes, I swear it's over). Without their love and support, this project would not have been possible.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
LIST OF TABLES	iv
CHAPTER	
I. INTRODUCTION	1
II. METHODS	11
III. RESULTS	26
IV. DISCUSSION	77
REFERENCES	104
APPENDIX A: EXTENDED LITERATURE REVIEW	110
APPENDIX B: INFORMED CONSENT	166
APPENDIX C: MATERIALS USED IN PARTICIPANT GROUPS	171
APPENDIX D: VERBAL INSTRUCTIONS	189

LIST OF TABLES

3.1	Number of Male and Female Participants in the Five Groups Conducted by each of the Three Experimenters	59
3.2	Interaction of Presence or Absence of Validity Information with Presence or Absence of Depression Information for each MMPI-2 Clinical Scale by Participant Gender	60
3.3	ANOVA Results for Comparisons of the Six Groups for each MMPI-2 Clinical Scale and the <u>DEP</u> Content Scale by Participant Gender	61
3.4	Clinical Scales and Depression Content Scale: Means, Standard Deviations, and Differences Between Groups (Male Participants Only)	62
3.5	Clinical Scales and Depression Content Scale: Means, Standard Deviations, and Differences Between Groups (Female Participants Only)	64
3.6	Interaction of Presence or Absence of Validity Information with Presence or Absence of Depression Information for each MMPI-2 Validity Scale or Index by Participant Gender	66
3.7	ANOVA Results for Comparisons of the Six Groups for each MMPI-2 Validity Scale and Index by Participant Gender	67
3.8	Validity Scales and Indices: Means, Standard Deviations, and Differences Between Groups (Male Participants Only)	68
3.9	Validity Scales and Indices: Means, Standard Deviations, and Differences Between Groups (Female Participants Only)	70
3.10	Optimal Cutting Scores, Percent Correctly Classified, and Average Hit Rates for MMPI-2 Validity Indices for Male Patients and Male Fakers in the Three Best Faking Groups	72
3.11	Optimal Cutting Scores, Percent Correctly Classified, and Average Hit Rates for MMPI-2 Validity Indices for Female Patients and Female Fakers in the Three Best Faking Groups	73

3.12	Percentage of Participants Correctly Classified using Cutting Scores for MMPI-2 Validity Indices that Represent the 90th Percentile in a Large Inpatient Sample	74
3.13	Percentage of Fakers Who Successfully Malingered by Gender and Group	75
3.14	Application of Criteria for Successful Faking to the Clinical Group: Percentage of Patients Correctly Identified as Non-Malingerers	76
4.1	Most Frequent Two Point MMPI-2 Profile Codetype Obtained by Each Group by Gender	102
A.1	Demographic and Methodological Characteristics and Mean D for 24 MMPI Studies of Malingering . .	157
A.2	Demographic and Methodological Characteristics and Mean D for 20 MMPI-2 Studies of Malingering .	162

CHAPTER I

INTRODUCTION

The introduction will provide an overview of the malingering issue and the rationale for the current research study. A more complete review of the previous studies can be found in the extended literature review (Chapter V).

The Diagnostic and Statistical Manual of Mental Disorders (4th ed; American Psychiatric Association, 1994) defines malingering as

the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives such as avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs (p. 683).

Faking mental illness on psychological evaluations given in forensic settings is particularly tempting for the potential malingerer because the rewards and benefits are so high. Persons who have been through a traumatic injury or even a minor injury have powerful incentives to either make up psychological symptoms, to exaggerate existing psychological disturbances, or to attribute existing symptoms to unrelated causes. Silvain's study (as cited in Lees-Haley & Dunn, 1994) indicated that workers' compensation or personal injury lawsuits often involve litigants with genuine disorders such as posttraumatic stress disorder (PTSD), major depression, generalized anxiety disorder, and mild brain injury). Although these psychiatric disorders can legitimately result from many types of trauma (Resnick, 1988), it is not hard to imagine that some people overreport or fabricate symptoms in hopes of gaining very large financial awards.

How do malingerers prepare to fake? Persons who fabricate mental disorders may have relatively accurate and specific information about the symptoms they need to report

(Wetter, Baer, Berry, Robison, & Sumpter 1993). This information can be obtained in several different ways. Some potential malingerers may have first or second-hand experience with specific disorders, either having grown up observing family members or friends who have a mental disorder or previously having had a mental disorder themselves. Most people have general knowledge of specific mental disorders through popular magazines, newspapers, television talk shows, and books (Lees-Haley & Dunn, 1994). Even television advertisements educate the public about mental illness, with some commercial psychiatric hospitals enumerating the symptoms of various psychological disorders in their advertisements. Lees-Haley (1992) reported that Los Angeles radio, television, magazine and flyer advertisements in many languages "solicit people to file personal injury and workers' compensation claims" (p. 682). These advertisements detail the symptoms that are necessary to report when filing a claim (Lees-Haley, 1992).

Besides learning about specific symptom information through the popular media, clients undergoing psychological evaluations may figure out what symptoms they are expected to report by the questions that attorneys or mental health professionals ask them (Ziskin & Faust 1988). Specifically, a study by Lees-Haley from 1985 (as cited in Ziskin & Faust, 1988) reported that lawyers' questioning "defines the issues and teaches a client the range of subjects likely to arise in the case." Scrupulous lawyers do not deliberately advise litigants to malingering psychopathology, but

the process of exploring possible psychological disorders can create a mental set in the litigant which encourages him to falsify or exaggerate symptoms . . . questioning may even cue him as to the kinds of symptoms that might be to his advantage. (Ziskin & Faust, 1988, p.851)

A recent survey of attorneys conducted by Wetter & Corrigan in 1994 (as cited in Baer, Wetter, & Berry, 1995) revealed

that a majority believed that attorneys should discuss upcoming psychological evaluations with their clients. Almost 50% of this sample of attorneys reported that attorneys should inform their clients about the presence of validity scales on psychological tests.

Besides obtaining symptom information passively or inadvertently, some sophisticated malingerers may actually prepare to fake by actively seeking out relevant symptom information from psychological textbooks and DSM-IV (Berry, 1995), or by learning about the test itself (Hare, 1985).

The MMPI and the MMPI-2 are probably the most widely used personality tests (Lubin, Larsen, & Matarazzo, 1984; Butcher, 1990). One of the reasons for the widespread use of the MMPI and the MMPI-2 is the test's ability to identify major patterns of mental disorder, such as depression and schizophrenia, while also detecting troublesome test taker approaches. In particular, both versions of the MMPI have been successful in detecting persons who "fake bad" or malingering (Berry, Baer, & Harris, 1991; Rogers, Sewell, & Salekin, 1994).

It is especially important to know more about the effective use of the MMPI-2 in detecting malingerers because the test is often used in settings where there is a real possibility of malingering. For example, the MMPI-2 is used in correctional facilities, in evaluating disability eligibility, and in other situations in which it might be advantageous for the test taker to feign mental illness.

There has been much research on the ability of the MMPI-2 to detect malingering. Much of the research has instructed participants to "fake bad" or, in other words, to attempt to appear maladjusted. More recently, MMPI-2 malingering studies have provided participants with information about a specific disorder and then asked them to attempt to fake that disorder. For example, participants have been instructed to fake specific mental disorders such

as PTSD and schizophrenia (Bagby et al., 1997; Rogers, Bagby, & Chakraborty, 1993; Sivec, Hilsenroth, & Lynn, 1995; Sivec, Lynn, & Garske, 1994; Wetter, Baer, Berry, & Reynolds, 1994; Wetter et al., 1993; Wetter & Deitsch, 1996). In these studies, participants were taught the symptoms of the disorder they were instructed to fake. This research has grown out of a concern that malingerers may prepare for taking the MMPI-2 by learning about specific mental disorders. Findings from these studies have suggested that information about specific disorders is relatively ineffective in producing successful malingering on the MMPI-2 (Rogers, 1997).

Other authors have proposed that some sophisticated malingerers may prepare to fake on the MMPI-2 by studying the test itself (Rogers et al., 1993; Lamb, Berry, Wetter, & Baer, 1994). Specifically, malingerers may attempt to learn how the MMPI-2 assesses malingering. Results from these studies suggest that MMPI-2 validity scale information is quite useful in helping faking participants elude detection on the MMPI-2 (Rogers, 1997).

Studying the effect of providing symptom information and validity scale information on malingering performance may help us understand two strategies that may be used by highly motivated malingerers. To date, no research study has assessed the impact of providing validity scale information, depression symptom information, or both types of information to participants who have been asked to fake depression on the MMPI-2. The only study specifically examining the malingering of depression on the MMPI-2 provided depression information to participants, but did not provide validity scale information (Bagby et al., 1997).

The goal of this study was to determine how well college students could malingering depression on the MMPI-2 when provided with two types of information: depression symptom information and MMPI-2 validity scale information.

Studying the ability of persons to malingering depression on the MMPI-2 is a relevant topic because major depression is very common in SSI disability claims (Griffin, Normington, May, and Glassmire, 1997; Reich, 1986) as well as in personal injury lawsuits (Silvain, 1992 as cited in Lees-Haley & Dunn, 1994).

Many individuals may recognize the advantages in either exaggerating existing sub-clinical depressive symptoms or in completely fabricating depression. Clearly, there are monetary and other incentives for a person to appear depressed, particularly in forensic settings and in situations where individuals are seeking disability status. In a recent study examining dissimulation among SSI disability claimants in Los Angeles, 100 claimants were tested (Griffin et al., 1997). Thirty-six percent of the claimants had complaints of depression, 43% for memory impairment, 34% for poor concentration, and 34% for auditory or visual hallucinations. In Griffin et al. (1997), malingering was assessed using a Composite Disability Malingering Index that was specifically created for the disability applicant population in Los Angeles. A cutting score was established that effectively differentiated between disabled nonmalingerers, instructed malingerers, and a population of disability applicants, some of whom were likely to be malingering and some of whom were not (Griffin et al., 1997). This cutting score found 19% of the disability claimants to be malingering. Clearly, malingering is an issue among SSI disability applicants and depression is one of the more frequent complaints of this population.

With the MMPI-2 being one of the most commonly used personality assessment measures, its effectiveness in the detection of malingering of all disorders must be thoroughly examined. Because it is unclear what kinds of information malingerers obtain or how they prepare to malingering, studying

the effect of providing depression symptom information and validity scale information on malingering performance is an important avenue of research. Clearly, not all malingerers are sophisticated or motivated enough to gather the types of extensive information that were provided to participants in this experiment.¹ But, by allowing participants to study validity scale and symptom information, they may have been as well-prepared as the most sophisticated and motivated real-life malingerer. Because the participants in this study were probably better prepared to fake on the MMPI-2 than the average malingerer, this study was a stringent test of the MMPI-2's ability to detect malingering.

Although current studies suggest that providing symptom information does not aid persons in malingering successfully on the MMPI-2, no studies have systematically examined the effect of validity information on the malingering of major depression. Depression and other non-psychotic disorders may be easier to fabricate on the MMPI than psychotic spectrum disorders. One reason for this expectation is that the public is better informed about depressive symptomatology than about psychotic disorders. Indeed, many untrained individuals can accurately endorse enough symptoms of depression to meet DSM-III-R criteria for Major Depression (Lees-Haley & Dunn, 1994). Furthermore, depression is a common disorder, with lifetime prevalence rates in community samples ranging from 10%-25% for women and 5%-12% for men (DSM-IV, 1994). Therefore, most people either have personal experience with depression or know a family member or friend who has had depression. Having this

¹Although the participants in this study were well-prepared to mangle, it is doubtful that they were as motivated as a genuine malingerer. It is virtually impossible to create a situation in an experiment that would approximate the kind of incentive and motivation that exist for real-life malingerers.

knowledge may be beneficial in faking depression on the MMPI-2.

In designing this study, an effort was made to improve upon the research methodology of previous studies. First, this study employed a scenario that was more appropriate and relevant to the college population than previous studies. Those studies have used scenarios asking students to imagine themselves in situations more suitable to a community sample. To maximize subject involvement and motivation, the present scenario was written with college students specifically in mind. Participants were asked to imagine feeling slightly depressed due to increasingly bad grades, fear of losing their academic scholarship, and financial difficulties. Participants were requested to imagine further that they discover that they may be eligible to receive monthly payments as well as deadline extensions and rescheduling of exams from professors if they are found to be severely depressed in a psychological evaluation including the MMPI-2. In contrast to some previous research which used relatively brief and uninvolved scenarios, care was taken in designing the present study to provide a longer scenario to enable participants to more easily imagine themselves in the situation suggested in the scenario.

Another difference between the present study and previous malingering research is the nature of the symptom information provided to participants. Although many studies have given participants symptom information based on DSM-III-R (1987) or DSM-IV (1994) criteria, few of them have mentioned the fact that persons do not have to have all the symptoms listed in order to receive the diagnosis. Unfortunately, this type of symptom information may unwittingly lead a person to exaggerate the disorder on the MMPI-2, thus leading to an inflated malingering detection rate. The current study presented the numbered DSM-IV criteria in a simplified and easy to understand manner. In

addition, the information clearly stated the number of symptoms needed for a diagnosis of major depression. Persons preparing to malingering depression may study lay literature or may not consult outside sources at all. If so, they would not have knowledge of the specific diagnostic criteria used in the DSM-IV. However, the sophisticated malingerer would have little difficulty obtaining access to psychology textbooks and to the DSM-IV. These sources spell out the criteria and clearly state that only a certain number of the criteria are necessary for diagnosis. Providing complete and understandable information about depression to participants--material which is readily available to the sophisticated malingerer--may make participants better able to fake depression on the MMPI-2 than many of their real-life counterparts. Assuming that symptom information helps malingerers fake more successfully and that a minority of real-life malingerers study the disorder they are attempting to fake, supplying participants with symptom information was a stringent test of the ability of the MMPI-2 to detect malingering.

This study also sought to improve upon past research by providing more understandable validity scale information, with an emphasis on strategies for eluding detection. In their information about validity scales and malingering strategies, two earlier studies presented not only tips for eluding the validity scales, but also validity scale names and descriptions of how the scales were developed. The validity information used in this study eliminated extraneous information such as scale names and scale development and history. Also, the use of technical terms such as "malingering" was avoided. The focus of the validity information in this study was on strategies for successful malingering. Although this type of simplified MMPI-2 validity information may not be available to the typical malingerer in the real world, these detection

strategies may be learned by the more intelligent and sophisticated malingerer through studying a detailed text on the MMPI-2. Providing faking participants with validity scale information was a rigorous test of the MMPI-2's efficacy in detecting malingering.

Measures were also taken to make sure that participants truly understood the information that was given to them. The participants were given time to absorb the information presented to them. In addition, they were given the opportunity to process the information by filling out a worksheet on the information and were quizzed on the information to test their comprehension of it. Therefore, persons who passed the quizzes (80% correct or better) were likely to have a firm grasp of information that may help them to fake depression on the MMPI-2.

A further improvement on existing methodology in this study involved employing a more extensive feedback questionnaire. The questionnaire was based partially on the debriefing materials used in an earlier study (Lamb et al., 1994). The questionnaire asked participants to evaluate their ability to stay engaged in the task and to stay in the role described in the scenario. In addition, the feedback questionnaire used in this dissertation assessed strategies used in attempting to successfully malingering using open-ended questions and checklists.

To date, no feedback questionnaires used in MMPI-2 malingering research have investigated participants' personal experience with mental disorders. Therefore, in an effort to assess the effect of personal experience with depression on malingering success, participants were asked if family members, friends, or they themselves had actually had depression. Although the existence of a clinical level depression cannot be inferred from such self-report information, relationships between MMPI-2 profiles and this information were examined to see if more systematic study of

this area is warranted. For example, if participants who reported having been depressed in the past were shown to be more successful at malingering depression on the MMPI-2 than participants who did not report previous episodes of depression, future research may want to investigate more systematically the effects of a past experience of a disorder on the successful malingering of that disorder.

A major limitation of MMPI and MMPI-2 research on malingering is the use of college students as participants in both experimental and control conditions and the lack of clinical comparison groups. Although most of the participants in this study were college students, one group was composed of clinically depressed inpatients. The use of a clinical comparison group allowed the current study to see if fakers could be differentiated not only from normal participants, but also from psychiatric patients.

Another limitation of some previous research in this area is small sample sizes. By using mostly college students as participants, a large sample was easily obtained for this study thus increasing statistical power. In addition, men and women were invited to participate in the study and were analyzed separately since it is well known the women and men manifest depression in different ways and because women and men respond differently to the MMPI-2 (Greene, 1991).

Based on previous research, it was predicted that validity information would enable faking participants in this experiment to elude detection by the MMPI-2 at a higher rate than the faking participants who did not received validity information. Since previous research has found that symptom information has not been helpful in malingering depression, it was predicted that depression information would not help participants successfully fake depression on the MMPI-2. For a detailed description of the hypotheses, see the "Hypotheses" subsection in Chapter II.

CHAPTER II

METHOD

Participants

Undergraduate participants

Male and female undergraduate students ($n = 568$) enrolled in an introductory psychology class at Texas Tech University participated in this research project in return for extra credit. Sign-up sheets containing a general description of the study and the length of time required for participation were posted.

Inpatient Samples

A clinical sample of depressed inpatients from three psychiatric facilities served as a comparison group ($n = 84$). The MMPI-2 responses from this group were used in an MMPI-2 study on the differential diagnosis of schizophrenia and major depression (Ben-Porath, Butcher, & Graham, 1991). The MMPI-2 profiles for these patients were collected between 1984 and 1986 and were derived from the larger MMPI-AX, the experimental form of the MMPI that was used in revising and restandardizing the MMPI. The MMPI-AX contained all the original MMPI items as well as 154 other experimental items. The MMPI-AX also contained all 567 MMPI-2 items (Ben-Porath et al., 1991).

The MMPI-2 data were collected at three different inpatient facilities, one in Ohio and two in Minnesota. Two facilities offered acute care. At these facilities, patients were tested in the early stages of their hospitalization. The third facility offered care for more chronic patients. At this facility, patients were tested at different phases of hospitalization and were offered \$4 to participate in the study (Ben-Porath et al., 1991).

Diagnoses were made by a psychiatrist using criteria listed in DSM-III (1980). At all three facilities the

diagnosis of depression was made independent of the MMPI-2 data (Ben-Porath et al., 1991). At the acute care facility in Minnesota, diagnoses were made by staff psychiatrists and psychiatric residents after consulting with a team consisting of a clinical psychologist, a psychiatric resident, psychology trainees, a psychiatric social worker, and psychiatric nurses. All the participants from this facility had the same diagnosis at discharge as at intake (Ben-Porath et al., 1991).

At the acute care facility in Ohio, diagnoses were made using the same process as described above for the acute care facility in Minnesota. Over 90% of the participants from this facility had the same diagnosis at discharge as at intake. Intake diagnoses were used because MMPI data were available in some cases to clinicians formulating discharge diagnoses (Ben-Porath et al., 1991).

At the state psychiatric hospital in Minnesota, the diagnosis was made by a psychiatrist following a clinical evaluation. Diagnoses were periodically revised and modified in consultation with a treatment team. The most recent diagnosis at the time the MMPI-AX was administered was used in this study (Ben-Porath et al., 1991).

All 84 patients had a primary diagnosis of major depression. Thirty-seven percent of the sample had a secondary diagnosis. Secondary diagnoses included 15 with chemical dependence/abuse, 2 with dysthymia, 6 with mixed personality disorder, 2 with antisocial personality disorder, 2 with borderline personality disorder, and 1 with histrionic personality disorder. Sixty-three percent of this patient sample did not have a secondary diagnosis (Ben-Porath et al., 1991).

Research Design

Two separate 2 X 2 factorial designs crossing the factors of presence and absence of depression information

with the presence or absence of validity scale information resulted in four groups of men and four groups of women who were asked to fake depression on the MMPI-2. Each group was given varying types of information and asked to fake depression. The first group was given information on symptoms of depression (for women, $n = 47$; for men, $n = 57$), the second group received information on MMPI-2 validity scales (for women, $n = 47$; for men, $n = 54$), the third group was given information on both validity scales and symptoms of depression (for women, $n = 76$; for men, $n = 77$), and the fourth group was given no information (for women, $n = 52$; for men, $n = 51$). The honest responding groups consisted of 52 female and 55 male participants instructed to take the MMPI-2 honestly.

Procedure

Participants were tested in small groups of no more than 20 individuals. Three different experimenters (all female) conducted the data collection sessions. Only participants belonging to the same experimental condition were tested together. Before beginning the experiment, participants in all conditions were required to read and sign an informed consent form (see Appendix B.1 for the consent form used for faking participants and Appendix B.2 for the consent form used for the honest respondents). All participants were assured of the anonymity of their responses.

Incentives

On the consent form, participants in the faking groups were told that cash prizes would be awarded to the top 5 fakers of depression. The prizes were \$75 for first place, \$30 for second place, \$20 for third place, \$15 for fourth

place, and \$10 for fifth place.² Participants were told that their individual responses would be kept strictly confidential and that all scores would be identified only by a code number. Participants who wanted to be eligible for a cash prize were encouraged by the experimenter to leave a name and phone number. Participants were informed that the principal investigator would keep a master list identifying participants by code number for the purpose of awarding the prizes, but once the prizes had been awarded, the master list would be destroyed.

Research procedures will be described separately for the groups receiving information, the no information group, and the honest responding group.

Procedure for Groups Receiving Information

Each participant received a packet of materials precoded with a participant number. The contents included a demographic information sheet; the experimental scenario; information on either depression, the MMPI-2 validity scales, or both types of information; a worksheet(s) for depression, validity scale information, or both types of information; and an MMPI-2 test booklet and answer sheet (see Appendices C.1-C.4).

After a brief explanation of the study by an experimenter (see Appendix D.1), participants were asked to fill out the demographic information sheet (see Appendix C.1). Then, a pre-recorded audio tape of the experimental scenario (see Appendix C.2) was played as participants read along silently. They were asked to imagine that they had experienced the situation described. After reading the scenario, participants in the depression symptom information

²An explanation of the criteria used to determine which participants were awarded cash prizes is provided in Chapter III in the subsection titled "Determination of successful faking."

group received a one-page handout describing the DSM-IV criteria for major depressive disorder (see Appendix C.3). Participants in the validity information group received a one-page handout describing the MMPI-2 validity scales and strategies used to "fake out" the MMPI-2.³ A pre-recorded audio tape of a person reading the appropriate information sheet was played and participants were requested to read along silently. Participants were then given 10 minutes to study the handout and to fill out a worksheet to help them learn and remember the depression symptoms (see Appendix C.4) or validity scale information. After 10 minutes, an answer key to the appropriate worksheet (see Appendix C.5) was distributed to each group member and they were asked to check their answers. In an effort to confirm that the information had been comprehended, participants were given a quiz on depression symptom information (see Appendix C.6) or validity scale information.

Participants in the group who received information on both depression symptoms and validity scales received both of the aforementioned information handouts, feedback sheets, and quizzes. Because of the large amount of material participants were required to learn in this experimental condition, they were given 10 minutes to study and complete the quiz for each type of information (20 minutes in all). The order of presentation of the information was counterbalanced across the groups that received both kinds of information.

After the quizzes were collected, participants completed the MMPI-2 using standard booklets and answer sheets. Participants were verbally reminded of their role (attempting to simulate depression) at the beginning of the

³Please see footnote regarding validity information in the subsection titled "Depression Symptom and Validity Scale Information."

MMPI-2 administration and again 20 minutes and 40 minutes into the test.

After completing the MMPI-2, participants were asked to fill out a feedback questionnaire (see Appendix C.7) which evaluated their level of motivation, use of faking strategies, and perceptions of their performance. Upon completion of the feedback questionnaire, students were given experimental credit for their participation, as well as a debriefing sheet (see Appendix C.8) thanking them for their involvement in the study, listing counseling opportunities on campus, and reminding them not to discuss the study with other introductory psychology students. To receive detailed information on the results of the study, participants were encouraged to contact the experimenter.

Procedure for the No Information Group

Each participant received a packet of materials precoded with a subject number. The contents included a demographic information sheet, the experimental scenario, and a MMPI-2 test booklet and answer sheet. After signing consent forms and filling out the demographic information sheets, participants were exposed to the experimental scenario in the same manner as the other three experimental groups. The scenario provided the instructions to fake depression. Next, participants took the MMPI-2, filled out the feedback questionnaire, and received experimental credit and a debriefing sheet using the same procedure as the other experimental groups.

Procedure for the Honest Responding Group

Each participant received a packet of materials precoded with a subject number. The contents included a demographic information sheet and a MMPI-2 test booklet and answer sheet. Using the standard MMPI-2 administration instructions, participants in the honest responding group

were asked to fill out the MMPI-2 in an honest and careful manner. After completing the MMPI-2, participants were asked to answer the quizzes on validity scale information and depression symptom information. Participants were asked to take the quizzes to gauge how well college students were aware of depression symptomatology and to provide a comparison with the groups who were given information sheets on depression and/or the validity scales. In addition, the honest participants were asked to fill out a brief version of the feedback questionnaire (see Appendix C.7 footnote). Participants then received experimental credit and a debriefing sheet describing the study, thanking them for their participation, and outlining counseling opportunities on campus.

Materials

Scenario

The scenario (see Appendix C.2) described a hypothetical situation involving a college student. Participants were asked to imagine feeling slightly depressed because of declining grades, fear of losing their academic scholarship, and financial difficulties. Each participant was requested to imagine further that he or she might be eligible to receive monthly payments as well as deadline extensions and rescheduling of exams from professors if he or she was found to be severely depressed in a psychological evaluation including the MMPI-2. The scenario asked each student to imagine exaggerating their mild symptoms of depression while responding to the MMPI-2 in hopes of receiving short-term financial and academic assistance. The scenario reminded participants of the benefits of malingering, specifically that being judged to be severely depressed could enable them to stay in school and to catch up on their bills. Participants were told that their goal was "to answer the MMPI-2 questions in such a way

as to make your minor depressive symptoms seem a lot worse than they actually are. Keep in mind that it is important to convince the test interpreter that you are severely depressed so that you can receive financial assistance and help from your professors. However, it will be obvious if you over-exaggerate, so it is very important that you answer the MMPI-2 in a believable manner."

Depression Symptom and Validity Scale Information

Depression information consisted of one single-spaced page presenting the numbered DSM-IV criteria in a simplified and easy to understand manner (see Appendix C.3). Symptoms such as significant changes in appetite and sleeping patterns, depressed mood, anhedonia, and difficulty concentrating were listed and explained. In addition, the number of symptoms needed for a diagnosis of major depression was clearly described. The information sheet also mentioned that the symptoms must interfere with daily life and represent a change from normal functioning.

Validity scale information consisted of one single-spaced page outlining six malingered strategies to use in order to avoid detection by the MMPI-2 validity scales.⁴

Depression and Validity Scale Information Worksheets and Quizzes

To increase the probability that participants processed as carefully as possible the information on depression and

⁴Due to the ethical concerns described in Chapter V, all descriptions of the validity information have been omitted from this document. Therefore, no examples of validity scale information are provided. In addition, the validity scale information sheet, validity scale worksheet, validity worksheet answers, and the validity scale quiz have been omitted from the document. Please see the end of Chapter V for instructions on how interested researchers may obtain this information.

the strategies for eluding detection, worksheets for the information and worksheet answer keys were provided (see Appendices C.4 and C.5). The depression worksheet contained five questions and the validity scale worksheet contained four. An item on the depression worksheet reads, "List as many symptoms of depression as you can--there are nine in all." A sample item from the validity scales worksheet asks participants to, "Name six strategies you can use to successfully fake a mental disorder on the MMPI-2." Feedback answer keys provided a clear answer to each worksheet item.

Participants were quizzed over the material to test their understanding of the information. Each quiz contained true-false and multiple choice questions (10 items in all). Items on the depression symptom information quiz (see Appendix C.6) assessed knowledge of DSM-IV criteria for major depression (e.g., "True or false: It is common for a person with depression to feel guilty."). Quiz questions about depression focused not only on symptom content, but also on which symptoms must be present to consider the diagnosis. The validity scales quiz focused mainly on the strategies for successful malingering. Participants who passed the quizzes (greater than 80% correct) were more likely to have a firm grasp of symptoms of depression, how to successfully malingering on the MMPI-2, or both.

Feedback Questionnaires

Based partially on the debriefing materials used by Lamb et al. (1994), the feedback questionnaire (see Appendix B.13) evaluated participants' perceived ability to stay engaged in the task and their perceptions of how well they were able to maintain the role of a depressed person. For example, participants were asked to rate on a 5-point Likert-type scale the following question: "How motivated were you to try to appear depressed on the MMPI-2?" In

addition, the feedback questionnaire assessed strategies used in attempting to malingering, using open-ended questions and checklists (e.g, "How did you attempt to appear depressed? In other words, what strategies did you use to fake out the MMPI-2?"). Participants were also asked if family members, friends, or themselves have actually had depression. In addition, they were asked the duration of the depressive episode, when it occurred, and if they had ever been hospitalized, attended therapy, or had an anti-depressant medication prescribed for them.

MMPI-2

Consisting of 567 true-false items, the MMPI-2 assesses a number of the major patterns of personality and emotional disorders (Butcher, Dahlstrom, Graham, Tellegen & Kaemmer, 1989). Designed for adults with at least an 8th grade reading level, the MMPI-2 generally takes 60 to 90 minutes to complete. T scores of 65 or over separate normal from abnormal scale elevations. The MMPI-2 has been found to have good test-retest reliability (1-week interval) with values ranging from .67 (scale 6) to .92 (scale 0) for men and from .58 (scale 6) to .91 (scale 0) for women (Butcher et al., 1989). Most reliability coefficients are .75 and over. Using Cronbach's alpha, internal consistency ranges from .34 (scale 6) to .85 (scales 7 and 8) for males and from .37 (scale 5) to .87 (scale 7) for females (Butcher et al., 1989). The MMPI-2 manual (Butcher et al., 1989) offers evidence for the validity of all of the MMPI-2 scales by showing their agreement with two measures, the Recent Life Events Survey and the Katz Adjustment Scale (KAS). However, the validity of the MMPI-2 relies mostly on the extensive validity information supporting the original MMPI (Dahlstrom, Welsh, & Dahlstrom, 1975).

Scale 2, comprised of 57 items, was designed to measure symptomatic depression (Greene, 1991). For the MMPI,

Dahlstrom et al. (1975) reported that Scale 2 had test-retest reliability coefficients ranging from .80 to .90 for intervals up to a month and ranging from .40 to .50 for intervals of one year or more. Greene (1991) reported that women tend to achieve higher scores on scale 2 than men. In terms of the validity of scale 2 on the MMPI-2, Boone (1995) correlated scores from the MMPI-2 scale 2 with scores from the Beck Depression Inventory, finding a validity coefficient of .56.

The depression content scale (DEP) has also been shown to have good concurrent and construct validity with correlation coefficients with other measures of depression, hopelessness, low self-esteem, and suicide ideation ranging from .58 to .66 (Boone, 1994). In addition, the DEP content scale has been found to significantly differentiate depressed from non-depressed psychiatric inpatients (Boone, 1994).

Hypotheses

Based on a review of the existing literature concerning malingering and the MMPI-2 (see Extended Literature Review, Chapter V), the following hypotheses are presented.

Hypothesis One

The depression information group, the depression and validity information group, and the no information group were predicted to achieve higher overall clinical scale elevations than the validity information group, the honest responding group, and the clinical group. In addition, the honest responding group was hypothesized to have significantly lower overall clinical scale elevations than the clinical group and the validity information group. The same pattern of results was expected for scale 2 and the DEP content scale. However, all four faking groups were expected to be able to elevate scale 2 and the DEP content scale (T score 65 or over).

Rationale for hypothesis one

The results predicted by hypothesis one are consistent with the pattern of results reported by other comparable MMPI-2 malingering studies. In general, these studies have shown that fakers provided with specific information achieve higher clinical scale elevations than honest responders (Lamb, et al. 1994; Rogers, et al. 1993; Sivec et al., 1994; Sivec et al., 1995; Wetter et al., 1994; Wetter & Deitsch, 1996). Thus, it was predicted that the faking groups would elevate not only scale 2 and DEP, but that they would elevate other scales that are not typically elevated by depressed patients. In general, these malingering studies on the MMPI-2 have shown that most participants are unsuccessful in faking a mental disorder on the MMPI-2, although groups provided with only validity scale information have a higher percentage of successful fakers. Using a similar design to the present study, Rogers et al. (1993) found that the no information group and the groups receiving symptom information (either alone or in conjunction with validity scale information) had significantly higher clinical scale elevations than the validity information group, the patient group, and the honest responder group. The validity information group and the patient group had similar elevations on clinical scales, although only the four clinical scales relevant to the malingered disorder were examined. In two studies employing patient comparison groups, malingerers achieved higher overall clinical scale elevations than bona fide patients (Rogers et al., 1993; Sivec et al., 1995), but one study found that the specific symptom information group differed significantly from patients on only four clinical scales.

The same pattern of results was predicted specifically for scale 2 and the DEP content scale (i.e., three faking groups would have significantly higher elevations on 2 and DEP than the validity information group, the clinical group,

and the honest responding group). The validity scale information group was hypothesized to have lower elevations on these two scales because it was believed that knowledge of the MMPI-2 validity scales in absence of the depression information would allow participants to be aware of items that tend to lead to extreme elevations, thus leading to moderate elevations on the clinical scales. All four faking groups were predicted to elevate scale 2 and the DEP scale because most malingering studies have shown that fakers tend to elevate most clinical scales, including scale 2. Since the DEP content scale is highly correlated with scale 2 (Greene, 1991), it was judged highly likely that the DEP content scale would be elevated if scale 2 was elevated. The only MMPI-2 malingering study with a fake depression group provided with depression symptom information found that this group elevated the 2 scale over a T score of 65 (Sivec et al., 1995).

Hypothesis Two

The validity information group was predicted to have significantly lower elevations on relevant validity scales and indices (F, F-K, Ds-r2, Ds-2, Fb, F(p), O-S) than the depression information group and the no information group. In addition, the validity information group was hypothesized to have significantly higher elevations on the validity scales and indices than the honest responding group. No differences were expected between the validity information group and the clinical group. There was no basis in the literature to predict significant differences between the the depression and validity information group and the other faking groups.

Rationale for hypothesis two

This hypothesis was based primarily on the findings of Rogers et al. (1993) and Lamb et al. (1994) who used a

similar design to the present study. Rogers et al. (1993) results showed that the validity information group scored significantly lower on F, F-K, Ds-r2, Fb than the schizophrenia symptom information group and the no information group, and significantly higher on these scales than the honest responding group. These relationships were not found for O-S. Rogers et al. (1993) also found that the validity information group did not differ from the schizophrenic patient group on F, Fb, and F-K, but had significantly higher elevations than the patient group on Ds-r2 and O-S. Consistent with Rogers et al. (1993), Lamb et al. (1994) reported that the validity information group achieved lower elevations on the validity scales than the other faking groups on F and Fb (the only validity scales examined), and lower elevations on these scales than the honest responding group. In the Lamb et al. (1994) study, participants were asked to malingering closed head injury. Previous research suggests that the depression information and validity information group would achieve validity scale elevations higher than the clinical group, the honest responding group, and the validity information group; however, often these differences have not been significant. The F(p) scale has never been tested in previous malingering studies using validity scale information as a variable, but has been shown to be significantly elevated by faking groups in other MMPI-2 malingering studies (Arbisi & Ben-Porath, 1995; Wetter & Deitsch, 1996; Wetter et al., 1994). It was expected that performance on the F(p) scale would follow the same pattern of results found for the majority of other validity scales and indices in previous malingering studies.

Hypothesis Three

It was predicted that cutting scores for the validity scales and indices (F, Fb, F-K, Ds-r2, Ds-2, F(p), and O-S) could be established that would lead to significant

differentiation between the clinical group and the faking groups, although this was predicted to be less likely to be true for the validity information group. This hypothesis was based on the assumption that most fakers would not be able to successfully malingering depression on the MMPI-2.

Rationale for hypothesis three

Because most MMPI-2 malingering studies providing specific information to faking groups have established cutting scores for most of the validity scales and indices that significantly differentiate between fakers and patients (Rogers et al., 1993; Sivec et al., 1995; Wetter et al., 1994, Wetter et al., 1993), it was assumed that the present study would also be able to establish such cutting scores. However, the study by Rogers et al. (1993), it was judged to be likely that the validity information group would be more difficult to differentiate from the clinical group using optimal cutting scores for the various validity scales and indices. Previous research offers no consensus as to which validity scales or indices are most successful in differentiating fakers from patients. For example, one study found that O-S and Ds-r2 were the most successful in differentiating fakers from patients (Rogers et al., 1993), while another study found that F(p) had the was most successful (Wetter et al., 1994). Overall, most validity scales and indices have been shown to be at least minimally effective in differentiating fakers from patients, especially when the fakers have been asked to "fake bad" or have been provided with specific symptom information.

CHAPTER III

RESULTS

The results for this study will be presented in the following order: (a) sample characteristics, (b) preliminary analyses and manipulation checks, (c) tests of the hypotheses, and (d) supplemental analyses. Tables referred to in this chapter are found at the end of the chapter.

Sample Characteristics

The description of the sample will be divided into three sections. The first section will describe the general demographic information about excluded student participants and the bases for exclusion of participants' data from this study. The second section will describe general demographic information about student participants whose data were retained. The third section will provide general demographic information about the clinical sample.

Excluded Student Participants

An attempt was made to assure that participants understood and complied with the instructions. Participants who were asked to fake depression were excluded from the analyses if they did not demonstrate mastery of the information presented on depression and/or validity scales (failed one or both of the quizzes with a score of less than 80% correct). If participants omitted over 15 MMPI-2 items, they were also excluded from the analysis. In addition, participants were excluded if they answered the MMPI-2 inconsistently or randomly as evidenced by a VRIN score equal to or greater than 13. Participants were also excluded if they achieved a VRIN score of 10 or greater, and they met both of the following two conditions: (1) indicated on the feedback questionnaire that they did not understand the instructions (score of 3 or less on a 5-point Likert

scale; see Appendix C.7) or that they were not motivated in faking depression while answering the MMPI-2 (3 or more), and (2) indicated on the feedback questionnaire that they answered 15% or more of the MMPI-2 items randomly. Using the above criteria, 78 men and 25 women were excluded from the data analysis, reducing the number of student participants from 568 to 465. The majority of excluded participants (69%) were eliminated because of high VRIN scores. Some of the same participants failed quizzes. Forty percent of excluded male and female participants came from the groups receiving both kinds of information. Participants were also excluded from the other conditions: a) depression information groups (for men, 25.5%; for women, 16%), b) validity information groups (for men, 7.5%; for women, 24%), c) no information groups (for men, 14%; for women, 12%), and d) honest responding groups (for men, 13%; for women, 8%).

The mean age of the excluded college participants was 19.0 (SD = .12) for men and 18.68 (SD = .37) for women. Ethnic identification for excluded men was 80% White, 6.5% Hispanic, 5% Asian, 5% African American, and 2.5% Native American. For excluded women, ethnic identification was 68% white, 20% Hispanic, and 12% African-American. The majority of excluded participants were freshmen in college (for men, 72%; for women, 76%).

Included Student Participants

The mean age of the included female college participants was 19.03 (SD = .21) and 19.47 (SD = .14) for male college participants. Ethnic identification was primarily White (80% of women, 84% of men), but there was a sizable Hispanic minority (11% of women, 8% of men). There were smaller numbers of African American (5% of women and 3% of men), Asian (2% of women and men), and Native American (1% of women and men) minorities. One percent of women and

men participants identified as "other." The majority of both male and female participants were freshmen in college (for women, 80%; for men, 63%). However, 13% of females and 21% of men were sophomores, 6% of women and 12% of men were juniors, and 1% of women and 4% of men were seniors or postgraduates.

There were no significant differences in education level or age when comparing excluded student participants to included college participants. For men, ethnic background was similar for both excluded and included student participants; however, for the women student participants, higher proportions of minorities (for Hispanics, 20%; for African Americans, 12%) were excluded from the analysis.

Clinical Sample

The mean age of female inpatients was 34.85 (SD = 11.07) and the mean age of male participants was 31.66 (SD = 11.34). Ethnic identification was primarily White (90% of women, 88% of men), with a small minority of African Americans (5% of women, 12% of men). Additionally, 2.5% of the female inpatients identified as Native American and 2.5% as "other." The average education level was 11.83 years for the women and 11.34 years for the men, with education levels ranging from grade 8 to 19 years of education for the women, and from grade 6 to 15 years for the men. Seventy percent of the female inpatients had a high school education or less and 30% had at least one year of college. For male inpatients, 63% had a high school education or less and 27% had at least one year of college. Two female inpatients were excluded from the study due to VRIN scores 13 or over leaving 40 female inpatients and 42 male inpatients in the clinical group.

Manipulation Checks and Preliminary Analyses

Order of Presentation for the Groups Receiving Both Kinds of Information

One of the faking groups received both validity scale and depression information. Since it was not known whether order of presentation would affect participant performance, presentation of the information was counterbalanced so that approximately half of the participants were exposed to the validity scale information first and half were exposed to the depression information first. Four sets of variables (quiz scores, the validity indices, the clinical scales and the Depression Content Scale, the feedback questionnaire items regarding understanding of instructions, motivation, and perceived performance) were contrasted using separate 2 X 2 MANOVAs with gender and the order of the presentation for the two counterbalanced groups as independent variables and the aforementioned scores as dependent variables. No significant main effects for order of presentation or significant interactions between gender and order of presentation were found for any of the dependent variables. Gender main effects (women scoring higher than men) were found on the clinical scales ($F [10, 99] = 4.30, p < .0005$) and on the validity scales ($F [7, 102] = 4.69, p < .0005$), but these differences are consistent with idea that men and women perform differently on the MMPI-2. Overall, these results suggest that the order of presentation did not affect the overall performance of participants who received both kinds of information.

Experimenter Differences

Since three different individuals (all female) were involved in collecting the data for this study, the possibility of experimenter differences was examined. Analyses were performed to determine if there were any systematic differences in participants' performance

depending on which experimenter ran the group and if there were any significant differences between experimenters in the number of male and female participants excluded from the study.

To examine participant performance, five 3 X 2 MANOVAs were performed with the experimenters and gender of the participants serving as the independent variables, and the validity scales, the clinical scales and the Depression Content Scale, quiz scores, and four feedback questionnaire items as the dependent variables. Each experimental group was examined separately for two reasons. First, there was an empty cell for one of the groups for one of the experimenters. Second, the number of participants as well as the proportion of men and women in the groups conducted by each experimenter varied greatly.

An attempt was made to make sure that each experimenter conducted experimental sessions for men and women in all experimental conditions (four faking conditions and the honest responding group). However, the proportions of men and women and the proportions of participants in the five experimental conditions differed among the three experimenters (see Table 3.1). One experimenter did not conduct an experimental session with any women in the No Information condition. Therefore, the 2 X 3 MANOVA could not be performed for the no information groups. Instead, five 2 X 2 MANOVAs with only two experimenters was performed for the no information group. In addition, a one-way MANOVA examining the male no information group with all three experimenters as the independent variable was performed for all the dependent variables.

On the quizzes, no multivariate or univariate main effects or interactions were found for any groups. However, on some of the other dependent variables, occasional significant multivariate main effects or interactions were found, but examination of the univariate main effects or

interactions did not reveal the same significant differences. For example, a multivariate main effect was found for the validity indices for the honest responding group ($F_{(14,166)} = 2.20, p < .01$), but none of the univariate interactions were significant for any validity index.

When the univariate Experimenter X Gender of Participant interactions were significant (depression information group on the Ds-2 scale, on the motivation question on the feedback questionnaire, on the DEP Content scale; validity information group on the Fp scale and scale 4; no information group on the perceived successful malinger performance on the feedback questionnaire and on Scale 2), the interactions had a consistent pattern. One experimenter tended to have male participants who obtained higher scores for the scales and group combinations mentioned above than the male participants run by the other two experimenters. The opposite pattern was seen for female participants: one experimenter tended to have female participants in her group that obtained lower scores than the female participants run by the other two experimenters. There is the possibility that this one experimenter behaved differently in her attitude or treatment of male and female participants in comparison to the other two experimenters thus biasing their data. However, it is impossible to know for sure if the differences observed were due to experimenter bias. For example, the differences may simply be spurious because some of the statistical effects were less powerful than some of the other effects found in this experiment. It is important to note that the pattern of the interaction was not pervasive across all the scales or all the groups. Rather, it was observed only on a very small percentage of individual scales and only for the experimental groups mentioned above.

A chi-square analysis was performed to determine if there were significant differences between experimenters in

the proportion of men and women excluded from the study. Results indicated that there were no significant differences (for men, $X^2 (2) = 1.51$, $p > .05$; for women, $X^2 (2) = .60$, $p > .05$).

Participant Compliance as Assessed by the Feedback Questionnaire

Included participants

An attempt was made to assess the level of participant motivation and compliance with experimental instructions by examining responses to several of the feedback questionnaire items, the VRIN scores, and quiz scores.

On the feedback questionnaire, faking participants were asked about their level of understanding of the instructions, their motivation level, and their perceptions of how successful their malingering performance was (see Appendix C.7). In addition, they were asked to estimate the percentage of MMPI-2 items that they answered randomly.

Analyses were performed to determine if there were any group or gender differences on these four feedback questionnaire items. Specifically, a $2 \times 2 \times 2$ MANOVA was performed with presence or absence of validity information, presence or absence of depression information, and the gender of the participants serving as the independent variables, and the four feedback questionnaire items as the dependent variables. No significant interactions were found, but a significant multivariate main effect was found for validity information ($F [4,356] = 3.32$; $p < .05$). In addition, an examination of the individual item ANOVAs revealed a main effect for validity information on the perceived malingering success question ($F [1,359] = 4.89$; $p < .05$) and for gender on the motivation question ($F [1,359] = 5.00$; $p < .05$).

Specifically, it was found that groups receiving validity scale information felt that they were less

successful in malingered depression ($M = 3.66$) than the groups that did not receive validity scale information ($M = 3.83$). Considering that the highest success rating on this item was 5.0, these means suggest that none of these groups felt very confident about their malingered performance.

An examination of the means by gender on the motivation question revealed that female fakers rated themselves significantly higher in motivation ($M = 4.32$) than the male fakers ($M = 4.12$), indicating that the female fakers perceived themselves to be more motivated in the experiment than male fakers (highest score possible was 5.0).

The honest responding group also filled out a feedback questionnaire, but, because they were not asked to fake depression, were only asked two of the four questions used in the aforementioned analysis. A separate 5 X 2 MANOVA with group membership (four faking groups and the honest responding group) and gender as independent variables and the two items (how well the instructions were understood and estimated percentage of items answered randomly) as dependent variables was performed. Multivariate and univariate main effects and interactions were nonsignificant. The mean score achieved by the groups was 1.27 indicating that participants reported that they understood the instructions that they were given. Across groups, the mean percentage of estimated random responses on the MMPI-2 was 13.4%.

To assess if there were any group or gender differences on level of inconsistent responding, a 2 X 6 ANOVA with gender of participants and group membership (all 6 groups included) as independent variables and the raw VRIN score as the dependent variable was performed. A significant interaction for gender and group membership was found for VRIN ($F [5, 527] = 2.37; p < .05$). In examining the interaction, women participants in the depression information group obtained significantly higher mean VRIN

scores than the male participants in the depression information group (for women, 6.02; for men, 5.11), but the women participants in the other three faking groups obtained lower mean VRIN scores than their male faking group counterparts.

Participant compliance and motivation was also assessed by examining performance on the validity and depression information quizzes. The three faking groups receiving information and the honest responding group took quizzes on validity scale information, on depression symptom information, or on both kinds of information. The honest responding participants received no preparation (i.e., information sheets) and took both quizzes after completing the MMPI-2. The honest responding groups were given the quizzes so that their scores could be compared to the faking groups to determine if the information given to the faking groups actually influenced their quiz performance.

To examine this issue, one-way ANOVAs were performed for each gender with group membership as the independent variable and one of the types of quizzes as the dependent variable. Faking students who were excluded from most of the other analyses based solely on their quiz scores were included in this analysis because students in the honest responding group could not be excluded from the analysis based on their quiz scores. However, student participants in the faking groups and the honest responding groups were excluded from this analysis if they achieved high VRIN or cannot say scores.

Three groups took the validity information quiz: the validity information group, the group receiving both kinds of information, and the honest responding group. For the validity information quiz, group membership was the independent variable and the validity information quiz was the dependent variable. For women, there were significant differences among groups ($F [2,162] = 199.35; p < .000001$).

Comparisons between groups using the Scheffe method revealed that the honest responding group obtained significantly lower quiz scores ($\underline{M} = 5.61$) than the two faking groups receiving validity scale information (for the validity information group, $\underline{M} = 9.33$; for the group receiving both kinds of information, $\underline{M} = 9.50$).⁵ The same pattern of results was observed for the men. There were significant differences among groups ($\underline{F} [2, 144] = 92.93$; $p < .0000001$), and the honest responding group earned significantly lower quiz scores ($\underline{M} = 6.13$) than the two faking groups receiving validity information (for the validity information group, $\underline{M} = 9.32$; for the group receiving both kinds of information, $\underline{M} = 9.11$).

Three groups took the depression information quiz: the depression information group, the group receiving both kinds of information, and the honest responding group. For the depression information quiz, group membership was the independent variable and the depression information quiz was the dependent variable. For women, there were significant differences among groups ($\underline{F} [2, 161] = 61.02$; $p < .0000001$). Comparisons between groups using the Scheffe method revealed that the honest responding group obtained significantly lower quiz scores ($\underline{M} = 7.76$) than the two faking groups receiving depression information (for the depression information group, $\underline{M} = 9.41$); for the group receiving both kinds of information, $\underline{M} = 9.40$). The same pattern of results was observed for the men. There were significant differences among groups ($\underline{F} [2, 137] = 52.61$; $p < .0000001$), and the honest responding group earned significantly lower quiz scores ($\underline{M} = 7.57$) than the two faking groups (for the depression information group, $\underline{M} = 9.26$; for the group receiving both kinds of information, $\underline{M} = 9.41$).

⁵ Quiz scores ranged from 1 to 10, with a 10 indicating a perfect score.

Overall, 22% of the men and 10% of the women in the honest responding groups passed the validity information quiz, whereas 95% of the men fakers and 97% of the women fakers passed the quiz. On the depression information quiz, 60% of the men and 63% of the women in the honest responding groups passed. For the faking groups taking the depression quiz, 97% of the men passed and 99% of the women passed. Clearly, the honest responding group scored much higher on the depression quiz than they did on the validity quiz. Many honest responders apparently were able to identify the symptoms of depression without having received information to study. However, in comparison to the faking participants who received information about depression, the honest responding groups had a significantly lower rate of success on the depression quiz. On the validity quiz, few honest respondents were able to pass, whereas almost all faking participants achieved passing scores. These results indicate that the information given to the faking groups was very important in participants' ability to pass the quizzes.

Excluded participants versus included participants

To compare excluded versus included participants on levels of compliance and motivation in this experiment, a 2 X 2 MANOVA was performed with participant gender and participant inclusion or exclusion as independent variables and the four feedback questionnaire items mentioned above as dependent variables. Only the four faking groups were included in this analysis because they were the only groups answering all four feedback questionnaire items. Results showed a significant multivariate main effect for inclusion/exclusion ($F [2,557] = 47.2, p < .0000001$). Univariate main effects for inclusion/exclusion were found for all four feedback questionnaire items. In comparison to the included participants, the excluded subjects rated

themselves lower on understanding of instructions, motivation level, and perceptions of malingering success, and rated themselves higher on estimated percentage of MMPI-2 items answered randomly.

As mentioned in the previous section, participants in the honest responding group were only asked two of the four feedback questionnaire items used in the aforementioned analysis. Therefore, a second analysis was performed which included the four faking groups and the honest responding groups. A separate 2 X 2 MANOVA was performed with participant inclusion/exclusion and gender as independent variables and two feedback questionnaire items (how well the instructions were understood and estimated percentage of items answered randomly) as dependent variables. As expected, there was a significant multivariate main effect for participant inclusion/exclusion ($F [2, 557] = 47.18$; $p < .0000001$) and significant univariate main effects for participant inclusion/exclusion on both feedback questionnaire items (for level of understanding feedback question, $F (1, 558) = 4.46$; $p < .005$; for percentage of items answered randomly, $F (1, 558) = 34588.14$; $p < .0000001$). Excluded participants reported answering a larger percentage of items randomly than the included participants (for excluded, $M = 36.9$; for included, $M = 13.5$). The direction of the main effects was identical to the previous 2 X 2 MANOVA that included only the four faking groups.

Tests of the Hypotheses

Statistical Analysis Overview

To successfully malingering depression on the MMPI-2, participants had to elevate clinical scales typically elevated by depressed persons and had to avoid elevating the validity scales in such a way as to reveal their malingering. Therefore, MMPI-2 profiles were examined first to determine if relevant clinical scales were elevated. On

the MMPI, scale 2 and the DEP content scale are commonly elevated by depressed women (Ben-Porath, Butcher, & Graham 1991; Wetzler, Marlowe, & Sanderson, 1994; Schipper, 1991). Participants were judged as successful malingerers of depression on the MMPI-2 if they scored 65 or over (T score) on scale 2 and DEP, if they scored significantly higher on scale 2 and DEP than the honest responding group, and if they did not excessively elevate relevant validity scales (F, F-K, Ds-r2, Ds-2, Fb, F(p), O-S).

Design

There were six groups of research participants in this study: four groups of college students who were attempting to malingering depression and two comparison groups (college students and depressed patients, neither of which were attempting to malingering). The profiles of men and women participants were analyzed separately. Therefore, in effect, there were 12 groups (6 groups of men and 6 groups of women). To assess the effect of depression information and validity scale information on the elevation of the nine clinical scales (omitting scale 5), the DEP content scale, and the validity indices, faking groups were compared to each other and to the honest responding group and the clinical group of depressed inpatients. Differences among the six groups were assessed in two steps. In the first step, the four faking groups were compared to each other in a 2 X 2 MANOVA, with depression information (presence vs. absence) and validity scale information (presence vs. absence) as independent variables. The nine clinical scales (excluding Scale 5) and the validity indices (F, F-K, Ds-r2, Ds-2, Fb, F(p), O-S) were used as dependent variables (using raw scores) in separate analyses. MANOVAs were used so that the clinical scales and the validity scales could be examined (the two sets of scales were examined separately) and so conclusions could be drawn about differences in the

profiles of the four groups. Including the 2 X 2 design allowed for assessment of whether depression information and validity scale information interacted significantly in influencing participants' scores on the clinical scales and on the validity indices.

Univariate 2 X 2 ANOVAs were used to determine the effect of the different types of information on participants' performance on each clinical scale, on the DEP scale, and on each validity index.

To compare clinical scale scores and validity indices scores among all six groups, a one-way MANOVA was performed with the clinical scales and the DEP Content Scale as dependent variables, and another one-way MANOVA was performed with the validity indices as dependent variables. The one-way MANOVA was followed by one-way ANOVAs for each clinical scale, for the DEP scale, and for the validity indices. Significant F ratios were followed by pairwise comparisons using the Scheffe method to determine the specific differences between groups on all scales and indices. The Scheffe test was selected because the family-wise error rate can be maintained at a specific, desired value ($p < .05$) regardless of the number of comparisons. Because successful fakers were expected to elevate both scale 2 and the DEP scale to a T score of 65 or over, group means on these scales were examined. Raw scores were used in all analyses (Greene, Gwin, & Staal, 1997).

Another goal of this study was to determine which validity scales were the most successful in differentiating the faking groups from the clinical group. For each validity scale, optimal cutting scores were established to distinguish the best faking groups from the clinical comparison group. To determine if optimal cutting scores could significantly differentiate fakers from patients, the percentage of each group that was correctly identified by each scale's cutting score was evaluated using the the

normal distribution, which provides a close approximation to the binomial distribution especially for large Ns.

Hypothesis One

The depression information group, the depression and validity information group, and the no information group were predicted to achieve higher overall clinical scale elevations than the validity information group, the honest responding group, and the clinical group. In addition, the honest responding group was hypothesized to have significantly lower overall clinical scale elevations than the clinical group and the validity information group. The same pattern of results was expected for scale 2 and the DEP content scale. However, all four faking groups were expected to be able to elevate scale 2 and the DEP content scale (T score 65 or over).

Multivariate results for the clinical scales

To examine the specific interaction expected (i.e., validity group scoring significantly lower than the other three faking groups), a 2 X 2 between-subjects MANOVA was performed with presence or absence of depression information and presence or absence of validity scale information as independent variables and with the MMPI-2 clinical scales as dependent variables. Only the four faking groups were included in the 2 X 2 MANOVA. The F ratios reported in all MANOVAs are approximate F values derived from the Pillai-Bartlett Trace statistic.

The 2 X 2 MANOVA for men using the clinical scales as dependent variables yielded significant main effects for depression information (F [9,159] = 3.05, p<.005) and validity scale information (F [9,159] = 4.37, p<.0001). The interaction between the two types of information was not significant (F [9,159] = .19, p>.05).

The same pattern was observed for the women on the 2 X 2 MANOVA. The main effects for depression information ($F [9,187] = 3.73, p < .005$) and validity scale information ($F [9,187] = 6.30, p < .000001$) were significant, but the interaction between the information was not significant ($F [9,187] = 1.71, p > .05$).

Individual clinical scale and DEP content scale results

Following the MANOVA, the univariate ANOVAs for each of the clinical scales were examined. For the men, the interaction between depression information and validity scale information was not significant for any individual clinical scale. For the females, even though the multivariate interaction was not significant, significant interactions were found for several individual scales (Scale 1, Scale 4, Scale 6, Scale 7, and Scale 8). For scales 1, 4, and 8, the pattern of the interaction was the same. Specifically, groups receiving both kinds of information or validity scale information in isolation achieved lower mean elevations than groups receiving depression information alone. The group receiving no information at all consistently obtained higher mean elevations on all four scales. (See Table 3.2 for specific F and p values for all the clinical scales.)

A 2 X 2 ANOVA including the four faking groups only was performed using the DEP content scale as a dependent variable. For men, the analysis yielded a significant main effect for validity scale information ($F [1,167] = 10.00, p < .005$). A nonsignificant main effect was found for depression information ($F [1,167] = 2.09, p > .05$) and for the interaction between the two types of information ($F [1,167] = .06, p > .05$).

A different pattern was observed for the women on the 2 X 2 ANOVA with the DEP content scale as a dependent

variable. The main effects for depression information ($F [1,195] = 4.03, p < .05$) and validity scale information ($F [1,195] = 22.12, p < .00001$) were significant, as was the interaction between the two types of information ($F [1,195] = 12.06, p < .001$). Specifically, the group receiving only validity information achieved a lower score than the other three faking groups (for the validity information group, $M = 22.51$; for the other three faking groups combined, $M = 27.36$).

Main effects will not be reported since the main purpose of the 2 X 2 MANOVAs and ANOVAs was to examine interactions of the two types of information. Results of the one-way analyses to follow provide detailed information on group differences.

Group Comparisons on clinical scale performance

To examine the group differences that were predicted in hypothesis one, a one-way MANOVA including all six groups was performed (i.e., the four faking groups plus the honest responding group and the clinical group). The clinical scale and the DEP content scale were the dependent variables. As expected, the one-way MANOVA was highly significant for men ($F [50,1111] = 7.12, p < .000001$) and for women ($F [50,1252] = 11.39, p < .000001$).

The one-way MANOVA was followed by one-way ANOVAs for each clinical scale. All the clinical scales and the DEP content scale were highly significant for men ($p < .00005$) and for women ($p < .000005$). (See Table 3.3 for F and p values.)

To determine specific differences between groups, significant F ratios were followed by pairwise comparisons using the Scheffe method (for men, see Table 3.4; for women, see Table 3.5).

For men, the honest responding group consistently scored significantly lower on the clinical scales and the DEP Content Scale in comparison with the faking groups and

the clinical group, with the exception of Scale 9 on which none of the groups differed from each other. For women, the honest responding group scored significantly lower on the clinical scales and the DEP content scale than all groups with the exception of Scales 1 and 9. On Scale 1, the female honest responding group did not differ from the clinical comparison group. On Scale 9, the female honest responding group did not differ from the clinical comparison group or any of the faking groups with the exception of the No Information group.

On four scales (1, 4, 6, Dep content scale), the male groups receiving validity scale information (both in isolation or in combination with depression information) achieved mean elevations that were similar to the clinical comparison group. This pattern was found only for Scale 4 for the women, although the group receiving validity scale information only did not differ significantly from the clinical comparison group on the DEP Content Scale.

On three scales (6, 8, and 0), all male faking groups attained significantly higher mean elevations than the clinical comparison group. This same pattern of results was found for women on Scales 1, 2, 7, 8, and 9.

On Scale 3, all male and female faking groups achieved elevations that did not differ significantly from the clinical comparison groups.

On some scales, the means of faking groups did not significantly differ from each other (for men, Scales 3, 4, 7, 0, and DEP Content Scale; for women, Scales 2, 3, and 0). On the remaining scales (Scales 1, 2, 6, 8, 9 for men; Scales 1, 4, 6, 7, 8, 9, and DEP Content Scale for women), the means of some faking groups were significantly different from each other. The general pattern observed among these scales was that the validity information groups (either alone or in combination with depression scale information) and the depression information groups achieved lower overall

means on the clinical scales and the DEP Content Scale than the no information group. In most cases, the validity information groups achieved the lowest means although they were not always significantly lower than the depression information group.

Conclusions regarding hypothesis one

It was predicted that a multivariate interaction and interactions for a majority of the clinical scales would be found. Overall, this prediction was not confirmed. Significant multivariate interactions for depression information and validity scale information were not found except for the female groups on the DEP Content scale (which was analyzed separately from the clinical scales in the first part of the analyses). Univariate interactions were found for a majority of the clinical scales for the women (Scales 1, 4, 6, 7, 8) despite the lack of a significant multivariate interaction. No multivariate or univariate interactions were found for men.

Overall, the comparisons between groups did not show the pattern predicted by Hypothesis 1 (i.e., the depression information group, the depression and validity information group, and the no information group were predicted to have significantly higher elevations on the DEP scale, scale 2, and on a majority of the clinical scales than all other groups). In many cases, the three faking groups mentioned in the hypothesis achieved similar elevations that were generally significantly higher than the honest responding groups, but not significantly higher on the majority of the clinical scales than the faking group receiving validity scale information only. On Scale 2, the predicted pattern of results was not found for men or for women. On the DEP Content Scale, the female groups showed the pattern of results predicted by hypothesis one, but the male groups did not.

It was predicted that the Validity Information group would achieve similar elevations as the clinical group on Scale 2, DEP, and on a majority of the clinical scales. It was further predicted that the elevations for the validity information group would be significantly higher than the elevations produced by the honest responding group. As predicted, the validity information groups scored significantly higher than the honest responding groups for both men and women. In addition the Validity Information groups scored similarly to the clinical group on a majority of scales (for men, 7 clinical scales; for women, 5 clinical scales). However, other faking groups (as mentioned above) also achieved similar elevations as the depressed impatient group, particularly the male group receiving both kinds of information. On five of the clinical scales, the four female faking groups achieved significantly higher elevations than the female clinical comparison group. On three of the clinical scales, the four male faking groups achieved significantly higher elevations than the male clinical comparison group.

To test the last part of hypothesis one, that each faking group would be able to elevate scale 2 and the DEP content scale to a T score of 65 or greater, the means of the four groups on scale 2 and the DEP content scale were examined. As predicted, the means of all four faking groups, both male and female, exceeded a T score of 65 for both Scale 2 and the DEP content scale (see Table 3.5).

Although isolated parts of Hypothesis one were confirmed, the overall results did not show the complete pattern of results that had been predicted (the female groups' performance on the DEP Content Scale was the only scale that showed the predicted pattern of results).

Hypothesis Two

The validity information group was predicted to have significantly lower elevations on relevant validity scales and indices (F, F-K, Ds-r2, Ds-2, Fb, F(p), O-S) than the depression information group and the no information group. In addition, the validity information group was hypothesized to have significantly higher elevations on the validity scales and indices than the honest responding group. No differences were expected between the validity information group and the clinical group. There was no basis in the literature to predict significant differences between the depression and validity information group and the other faking groups.

Multivariate results for the validity indices

A 2 X 2 MANOVA was performed with presence or absence of depression information and presence or absence of validity scale information as independent variables and the validity indices (F, F-K, Ds-r2, Ds-2, Fb, F(p), O-S) as dependent variables. Only the four faking groups were included in this analysis.

For the men, the main effects for depression information (F [7,161] = 2.36, p<.05) and validity scale information (F [7,161] = 5.48, p<.00005) were significant, as was the interaction between these types of information (F [7,161] = 2.45, p<.05).

For women, the main effects for depression information (F [7,189] = 4.57, p<.0001) and validity scale information (F [7,189] = 6.67, p<.01) were significant, but the interaction between the information was not significant (F [9,187] = 1.98, p>.05).

Individual validity index results for hypothesis two

Following the MANOVA, the univariate 2 X 2 ANOVAs for each of the validity scales were examined. For the men, the

interaction between depression information and validity scale information was significant only for the Fb scale ($F [1,167] = 4.19; p < .05$). Specifically, the no information group scored much higher than the other three faking groups on Fb. For the females, even though the multivariate interaction was not significant, significant interactions were found for all of the individual validity indices. For 6 of the 7 validity indices (F, Fb, Fp, Ds-2r, Ds-2, and F-K), the pattern of the interaction was the same. Specifically, the group receiving validity scale information obtained the lowest mean elevations, followed closely by the group receiving both kinds of information. The group receiving depression information only or no information obtained higher means than the other two faking groups. (See Table 3.6 for specific F and p values for all the validity indices.)

Group Comparisons on validity indices performance

To examine the group differences predicted in hypothesis two, a one-way MANOVA including all six groups was performed (i.e., the four faking groups plus the honest responding group and the clinical group). As expected, the one-way MANOVA was highly significant for men ($F [35,1037] = 8.11, p < .000001$) and for women ($F [35,1167] = 10.6, p < .000001$).

The one-way MANOVA was followed by one-way ANOVAs for each validity scale or index. There were significant differences among the six groups for all the validity indices, both for men ($p < .00001$) and for women ($p < .000001$) (see Table 3.7).

To determine specific differences between groups, significant F ratios were followed by pairwise comparisons using the Scheffe method (for men, see Table 3.8; for women, see Table 3.9 for detailed results). For men, all the faking groups achieved significantly higher mean scores on

all the validity indices than the clinical comparison group. This was also true for the women faking groups, with the exception of Fb and Fp. On these two indices, the Validity Information group did not differ from the clinical comparison group. On a majority of the validity indices, the validity information group did not differ from two other faking groups: the depression information group and the group receiving both kinds of information (for men, 5 indices; for women, 6 indices). For the women, the no information group scored significantly higher than the three faking groups on six out of seven validity indices (all except Fb). Although the male no information group scored higher than the other faking groups on all indices, this group was never significantly higher than all of the three other faking groups. However, this group was often significantly higher than the validity information group and the group receiving both kinds of information. The no information group did not differ from the depression information group on any validity index. For both men and women, all the faking groups scored significantly higher than the honest responding group. For most validity indices, the honest responding group did not differ from the clinical group either for men or for women.

Conclusions regarding hypothesis two

It was predicted that the validity information group would obtain significantly lower scores on most of the validity scales than the depression information group and the no information group, and that the validity information group would obtain significantly higher elevations on most of the validity scales than the honest responding group.

For the men, hypothesis two was partially confirmed. The validity information group obtained significantly higher elevations on all seven validity scales than the honest responding group. The validity group also obtained

significantly lower elevations than the no information group on six validity indices, but did not attain significantly lower means than the depression information group on six validity indices.

For the women, hypothesis two was also partially confirmed. On six out of seven validity indices, the validity information group obtained significantly higher elevations than the honest responding group. On all seven indices, the female validity information group had significantly lower elevations than the No Information group. However, the validity information group did not score significantly lower than the depression information group on six indices or the group receiving both kinds of information on all seven indices.

Although no differences on validity indices were expected between the validity group and the clinical group, the male validity group scored significantly higher than the clinical group on all validity indices. The female validity group also scored significantly higher than the clinical group on all but two validity indices (Fb and Fp).

Hypothesis Three

It was predicted that cutting scores for the validity scales and indices (F, Fb, F-K, Ds-r2, Ds-2, F(p), and O-S) could be established that would lead to significant differentiation between the clinical group and the faking groups, although this was predicted to be less likely to be true for the validity information group. This hypothesis was based on the assumption that most fakers would not be able to successfully malingering depression on the MMPI-2.

Optimal cutting scores

To test hypothesis three, a frequency distribution of scores for each validity scale or index (F, Fb, F-K, Ds-r2, Ds-2, F(p), and O-S) was examined to determine optimal

cutting scores--scores that best identified participants as being from the clinical or faking groups. Since the focus of this experiment was to examine how well a sophisticated malingerer supplied with pertinent information could fake depression on the MMPI-2, the most successful faking groups (those groups receiving validity scale information, depression information, or both types of information) were combined to form one group. They were compared to the clinical depression group to establish optimal cutting scores. Cutting scores were figured separately for men and women for seven validity indices. Tables 3.10 and 3.11 list the optimal cutting scores, the percentage of fakers and patients correctly classified by that cutting score, and an average hit rate which was figured by averaging the correct classification percentages of fakers and patients.

For men, the average hit rate ranged from 72 to 81%, with the F scale and the O-S index demonstrating the lowest average hit rates and the Revised Dissimulation scale achieving the highest average hit rate. The highest percentage of patients correctly classified was 79% as achieved by the Revised Dissimulation scale. The Obvious minus Subtle Index correctly classified the highest percentage of fakers (88%), but at the expense misclassifying 43% of the clinical group. The optimal cutting scores were able to correctly classify 71 to 88% of the 3 best faking groups combined, however, often at the expense of poor classification rates for the inpatient group (e.g., 65% [women] and 67% [men] correctly classified by an F scale cutting score of 14).

The pattern of results was somewhat different for the women. The Revised Dissimulation scale and the F scale cutting scores achieved the highest average hit rate (74%) and the Fp scale achieved the lowest average hit rate (68%). Overall, using these optimal cutting scores, the female fakers had higher correct classification rates than the men,

but the female patient group had a higher misclassification rate than the male patient group.

Using the the normal distribution, which provides a close approximation to the binomial distribution especially for large Ns, the classification accuracy percentages produced by the optimal cutting scores were evaluated for statistical significance. Using this method, the overall percentage of correct classification (average hit rate) identified by the cutting scores for the seven validity indices presented above were found to be statistically significant; in other words, the classification percentages were significantly greater than chance (i.e., 50%).

For men, the overall classification accuracy (combining fakers and patients) was statistically significant for the optimal cutting score for each validity scale or index (all p s < .0005). For fakers, at least 77 individuals (out of the total 131) needed to be correctly identified for the classification accuracy to be statistically significant (p < .05), so each validity index achieved that standard. For patients, at least 28 individuals (out of 42) needed to be correctly classified to reach statistical significance. Therefore, the classification of male patients by the optimal O-S cutting score was not statistically significant (z' = .93, p > .05). All other validity scales and indices reached statistical significance for male patients.

For women, the overall classification accuracy (combining fakers and patients) was statistically significant for each validity scale or index (all p s < .0005). For fakers, at least 88 individuals (out of 150) needed to be correctly classified to reach p < .05, which was true for each validity scale or index. However, only the F(p) scale met the standard that at least 27 individuals among the 40 patients needed to be correctly classified to reach statistical significance. The z' values for women

patients are as follows: 1.90 for F, 1.58 for Fb and F-K, .32 for Ds-2, .95 for Ds-r2, and .63 for O-S.

Discriminant function analysis

In addition to the cutting scores, a discriminant function analysis was performed looking at the same groups that were used for the cutting scores. To determine if a combination of the validity indices could improve upon the best cutting score, a discriminant function analysis was performed comparing the three best faking groups combined and the clinical comparison group.

A forward stepwise model was used in the discriminant function analysis (Statistica, 1994). Validity indices were added into the discriminant function model one by one. At each step, the validity index that made the most significant contribution to the discrimination between groups was added to the model.

For men and women, only three validity indices (Ds2r, Fp, and Fb) were added to the model. For women, the discriminant function correctly identified 65% of the clinical group and 86% of the combined best faking groups, for a 81.6% average hit rate for correct classification. This classification by discriminant function was an improvement on the average hit rate (74%) produced by using the optimal cutting score of 14 for the Ds-r2 scale, which was the best single measure. In addition, a larger proportion of the clinical group was correctly classified using the discriminant function (65%) in comparison with the Ds-2r cutting score (58%).

The discriminant function for the men with Ds-2r, Fp, and Fb in the model accurately classified 79% of the clinical group and 78% of the best faking groups producing an overall correct classification rate of 78%. In contrast to the women, the discriminant function for men did not improve upon the single optimal cutting score for Ds-r2. A

Ds-r2 optimal cutting score of 14 accurately classified 79% of clinical group and 82% of the faking group with an overall correct classification rate of 81%. For classification accuracy using validity indices cutting scores that have been suggested for use in correctly identifying 90% of an inpatient population, see Supplementary Analyses section.

Conclusions regarding hypothesis three

Hypothesis three predicted that cutting scores for the validity scales and indices (F, Fb, F-K, Ds-r2, Ds-2, F(p), and O-S) could be established that would lead to significant differentiation between the clinical group and the faking groups, although this was predicted to be less likely to be true for the validity information group. Hypothesis three was partially confirmed because cutting scores were established that led to significant differentiation between the depressed faking group and the best three faking groups. Because the validity information group did not perform significantly better across the board than the two other groups receiving information, all three groups were combined and compared to the clinical group to determine optimal cutting scores. Due to the similar scores obtained by three of the faking groups, the prediction that it would be more difficult to obtain cutting scores for the validity information group than the other faking groups was not examined.

Supplementary Analyses

The supplementary analyses will include the following sections: (a) correct classification rates of groups using normative cutting scores, (b) percentages of each group who successfully faked depression according to two different sets of criteria, and (c) an examination of the effect of

having experienced depression or having known a person with depression on malingering performance.

Correct classification of participants using
normative validity indices cutting scores

An alternative method of deriving cutting scores was used to assess participants' performance on the MMPI-2 validity indices. To determine how groups performed on the individual validity scales, cutting scores established using an independent inpatient sample were applied to all the groups. The cutting scores used were recommended for use in inpatient settings and correctly classified 90% of a large inpatient sample (Greene, 1997).⁶ In choosing the 90th percentile cutting scores, it was reasoned that correct identification of individuals with bona fide psychopathology (i.e., not mistakenly identifying a genuinely mentally ill person as a malingerer) was more important than correct identification of malingerers. Clearly, misidentifying a genuinely psychologically distressed person as a malingerer could result in serious consequences for the patient and the inpatient facility that made the incorrect diagnosis. However, because the need to identify malingerers is also important, a more extreme percentile cutoff (i.e., 95th percentile) was avoided in favor of the 90th percentile. Table 3.12 lists the percentage of each group correctly classified by the cutting score, reported separately for men and women. There was no recommended cutting score for Ds2,

⁶ When using the term "correctly classified" with these cutting scores, the assumption is being made that the score at the 90th percentile represents a cutting point such that scores above that point are likely to be invalid and scores below that point are likely to be valid. However, using the term "correctly classified" is somewhat misleading because the cutting scores used do not constitute an independent criterion of what is "correct." There is a possibility that scores above the 90th percentile could be obtained by individuals with severe mental illness.

therefore only six validity indices are listed in Table 3.12.

The groups receiving validity scale information, whether alone or in combination with depression information, were more successful in eluding detection by the validity indices. For example, the highest rate of correct classification of malingering in the male validity information group was 38% as achieved by the Fp scale. For the female validity information group, the Fp scale was able to correctly identify 49% as malingerers.

The groups receiving no information were much less successful than the other three faking groups. The highest rate of correct classification of malingering in the male no information group was 75% as achieved by the Fp and the Fb scale. For the female no information group, the Fp scale produced the highest classification rate (82%).

The Obvious-Subtle index did the poorest job of classifying the fakers with successful identification of malingering ranging from 8% (female group receiving both kinds of information) to 40% (male no information group).

On average, an Fp scale cutting score of greater than 5 was the most successful in detecting malingering among the male participants with correct classification rates ranging from 38% (validity information group) to 75% (no information group). The Fp and the Fb scales were the most successful in detecting malingering among the female participants with correct classification rates for Fp ranging from 39% (group receiving both kinds of information) to 82% (no information group) and classification rates for Fb ranging from 34% (validity information group) to 76% (no information group).

Determination of successful faking

To determine how many fakers were able to successfully malingering depression on the MMPI-2, two different methods were used.

Method one

Participants were judged as successful malingerers of depression on the MMPI-2 if they obtained a T score of 65 or over on both Scale 2 and DEP Content scale and if they did not elevate any of the six validity scales and indices using the cutting scores reviewed in the previous section (i.e., scores correctly classifying 90% of a sample of inpatients).⁷ Table 3.13 summarizes the percentage of persons in each faking group that were able to meet this criteria.

The pattern of results was different for men and women. For men, the most successful faking group was the group receiving only validity scale information (42% were able to elude detection), but, for the women fakers, the group receiving both kinds of information was the most successful (34% were able to elude detection). For both men and women fakers, the no information group was the least successful in malingered depression on the MMPI-2 (for men, 5%; for women; 8%).

Applying this "successful faking" criteria to the clinical group revealed high misclassification rates (for male inpatients, 48%; for female inpatients, 72%). The low rate of correct classification can be accounted for by the fact that many individuals in the clinical group did not elevate both Scale 2 and the DEP Content scale.

Method one "successful faking" criteria were not applied to the honest responding group. Applying these

⁷The five most successful fakers (winners of cash prizes) were determined using a variation of Method one. Specifically, scale 2 and the DEP content scale had to be 65 or over (T-score) and scores on the validity scales had to be lower than the validity scale indices scores representing the 75th percentile (Greene, 1997). Out of the remaining 20 successful fakers, the five participants with the lowest overall scores on the six validity indices were awarded the cash prizes.

criteria to the honest responding group is inappropriate because honest test takers in an experimental situation do not have an incentive to malingering.

Method two

A second set of criteria to judge successful malingering performance was created that reflected better the actual pattern of MMPI-2 results among the clinical group. Not all depressed persons attain elevations on Scale 2. Furthermore, a diagnosis of depression is not dependent on an elevation on scale 2. In fact, in everyday practice, a whole range of MMPI-2 profile patterns are seen in depressed individuals. An examination of the mean MMPI-2 profiles of the clinical group on the clinical scales revealed elevated scores for scales 8, 7, 2, 4, and 6 for men and scales 6, 8, 2, 7, and 4 for women.

Using this information about the actual MMPI-2 results of depressed inpatients, participants were judged as successful malingerers of depression on the MMPI-2 if they obtained a T score of 65 or over on any of the following scales: Scales 2, 4, 7, 8, 6, and DEP Content scale and if they did not elevate any of the six validity scales and indices using the cutting scores reviewed in the previous section. Table 3.13 summarizes the percentage of persons in each faking group that were able to meet this criteria.

The pattern of results using the second criteria for faking was the same as the pattern seen for the original criteria for faking. For men, the most successful faking group was the group receiving only validity scale information (50% were able to elude detection), but, for the women fakers, the group receiving both kinds of information was the most successful (38% were able to elude detection). Using the second criteria, more fakers were able to elude detection.

Applying the second criteria for successful faking to the clinical group revealed lower misclassification rates (for male inpatients, 29%; for female inpatients, 37%) than the rates seen using the original criteria (see Table 3.14).

Effect of personal experience of depression on malingering performance

After reading the simplified DSM-IV criteria for depression, all student participants were asked via the feedback questionnaire if they were currently depressed, had been depressed in the past, or had never been depressed. Those who indicated that they had been depressed were asked if they had been diagnosed with depression. Furthermore, participants were asked if they had ever known someone who was depressed. The effect of personal experience with depression on malingering performance was assessed using chi-square analyses. Results indicated that personal experience with depression (either having experienced depression in the past or having known someone who experienced depression) did not enable participants to elude detection at significantly higher rates than participants who did not have personal experience with depression.⁸ For example, successful and unsuccessful male fakers did not differ as a result of having experienced depression in the past or currently ($\chi^2 (1) = 2.46, p > .05$).

⁸ Method one criteria were used to distinguish successful from unsuccessful fakers.

Table 3.1

Number of Male and Female Participants in the Five Groups Conducted by each of the Three Experimenters.

Group	Experimenters		
	#1	#2	#3
Depression			
Men	27 (9)	27 (10)	3 (1)
Women	18 (2)	15 (2)	14 (0)
Validity			
Men	17 (2)	34 (3)	3 (1)
Women	9 (1)	27 (3)	11 (2)
Both			
Men	37 (15)	22 (13)	18 (3)
Women	39 (6)	18 (2)	19 (2)
No Info			
Men	11 (1)	13 (4)	27 (6)
Women	28 (2)	24 (2)	0 (0)
Honest			
Men	19 (5)	28 (4)	8 (1)
Women	8 (0)	21 (1)	23 (1)

Note. Numbers not in parentheses include all participants prior to the application of exclusion criteria. Numbers in parentheses refer to participants that were excluded from the study. Exclusionary criteria are listed in the "Sample Characteristics" section of the results. Depression = fakers receiving information about symptoms of depression, validity = fakers receiving information about MMPI-2 validity indices, both = fakers receiving information about symptoms of depression and information about MMPI-2 validity indices, no info = fakers receiving no information; honest = participants asked to take the MMPI-2 in an honest manner.

Table 3.2

Interaction of Presence or Absence of Validity Information with Presence or Absence of Depression Information for each MMPI-2 Clinical Scale by Participant Gender.

	Men	Women
Scale	<u>F</u> (1,167)	<u>F</u> (1, 195)
<u>1</u>	.69	5.35*
<u>2</u>	.04	.11
<u>3</u>	.05	.53
<u>4</u>	.00	4.68*
<u>6</u>	.53	3.79
<u>7</u>	.06	10.87**
<u>8</u>	.08	6.51*
<u>9</u>	.04	2.16
<u>0</u>	.01	.82

Note. Analyses were performed using raw scores for all scales.

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

Table 3.3

ANOVA Results for Comparisons of the Six Groups for each MMPI-2 Clinical Scale and the DEP Content Scale by Participant Gender.

	Men	Women
Scale	<u>F</u> (5,252)	<u>F</u> (5,283)
<u>1</u>	23.78*	35.05*
<u>2</u>	53.66*	75.72*
<u>3</u>	20.71*	18.16*
<u>4</u>	34.05*	51.87*
<u>6</u>	24.10*	36.09*
<u>7</u>	38.98*	59.68*
<u>8</u>	40.01*	54.68*
<u>9</u>	5.97***	7.05**
<u>0</u>	43.10*	96.90*
<u>DEP</u>	48.67	88.24*

Note. Analyses were performed using raw scores for all scales.

* $p < .0000001$, ** $p < .00001$, *** $p < .0001$

Table 3.4

Clinical Scales and Depression Content Scale: Means, Standard Deviations, and Differences Between Groups (Male Participants Only).

Scale	Faking Groups				Non-Faking Groups	
	Dep+Vala	Depb	Valc	No Infod	Honeste	Clinicalf
<u>1</u> (Hs)	13.5ad (4.8)	17.3ab (6.1)	14.6ad (6.9)	19.9b (5.5)	6.9c (3.9)	12.4d (7.8)
<u>2</u> (D)	36.7abe (7.3)	38.4b (4.0)	33.5ae (6.6)	35.7abe (5.1)	19.2c (5.1)	32.2e (8.2)
<u>3</u> (Hy)	27.2a (5.3)	29.2a (5.5)	27.4a (6.0)	29.8a (4.2)	19.6b (4.5)	28.8a (6.9)
<u>4</u> (Pd)	29.6ac (4.5)	32.1a (4.7)	29.3ac (6.2)	31.9a (5.6)	18.7b (5.4)	27.5c (6.6)
<u>6</u> (Pa)	16.5ab (4.3)	19.6bc (4.8)	16.2a (4.6)	20.3c (4.2)	10.8d (3.2)	16.0a (5.5)
<u>7</u> (Pt)	32.8a (6.4)	36.4a (5.7)	31.7a (7.2)	35.8a (5.3)	17.6b (7.8)	25.9c (11.0)

Table 3.4
Continued.

Scale	Faking Groups			Non-Faking Groups		
	Dep+Val ^a	Dep ^b	Val ^c	No Info ^d	Honest ^e	Clinical ^f
<u>8</u> (Sc)	37.1a (10.9)	45.9b (11.6)	37.4a (10.8)	47.2b (9.4)	18.36c (9.3)	28.8d (14.4)
<u>9</u> (Ma)	19.2a (4.6)	21.6ab (4.5)	20.3ab (3.8)	23.0bc (4.1)	21.2ac (3.8)	18.6a (4.7)
<u>0</u> (Si)	48.3a (10.1)	51.2a (8.0)	47.4a (10.2)	50.6a (6.6)	26.9b (9.1)	39.3c (11.2)
<u>DEP</u>	24.5a (6.2)	27.0a (3.8)	23.0ac (7.1)	26.0a (3.8)	8.5b (6.8)	19.8c (8.8)

Note. All scales are reported as raw scores. Mean scores with common subscripts are not significantly different (alpha = .05) using the Scheffe procedure. Standard deviations are enclosed in parentheses. Dep + Val Info = fakers receiving information about symptoms of depression and information about MMPI-2 validity indices; Dep Info = fakers receiving information about symptoms of depression and information about symptoms of depression; Val Info = fakers receiving information about MMPI-2 validity indices; No Info = fakers who were not given symptom or validity scale information. ^an = 66. ^bn = 43. ^cn = 41. ^dn = 49. ^en = 50. ^fn = 40.

Table 3.5

Clinical Scales and Depression Content Scale: Means, Standard Deviations, and Differences Between Groups (Female Participants Only).

Scale	Faking Groups				Non-Faking Groups	
	Dep+Vala	Dep ^b	Val ^c	No Infod	Honeste	Clinical ^f
<u>1</u> (Hs)	17.6a (5.9)	18.4ab (6.5)	17.7a (6.3)	22.6b (6.1)	8.0c (4.9)	11.9c (7.1)
<u>2</u> (D)	39.9a (5.0)	40.6a (6.6)	36.6a (7.3)	37.9a (4.9)	21.1b (5.6)	30.9c (7.2)
<u>3</u> (Hy)	29.4a (5.1)	31.7a (6.4)	28.8a (6.6)	32.2a (5.3)	22.4b (4.7)	28.5a (6.6)
<u>4</u> (Pd)	28.9ad (4.5)	31.7ab (5.2)	28.3ad (6.6)	34.4b (4.8)	18.3c (5.5)	25.7d (6.1)
<u>6</u> (Pa)	17.0a (4.2)	19.6ab (4.6)	17.1a (4.3)	22.1b (4.3)	11.0c (3.3)	16.8a (5.4)
<u>7</u> (Pt)	35.5ab (4.5)	35.7ab (6.6)	33.1a (7.0)	38.2b (4.5)	17.9c (8.0)	26.8d (10.7)

Table 3.5

Continued.

Scale	Faking Groups				Non-Faking Groups	
	Dep+Val ^a	Dep ^b	Val ^c	No Infod	Honest ^e	Clinical ^f
<u>8</u> (Sc)	38.5a (9.8)	44.3ab (12.5)	38.2a (12.5)	52.4b (11.5)	16.6c (10.2)	28.3d (14.9)
<u>9</u> (Ma)	18.2a (5.3)	19.6a (5.7)	19.6a (4.4)	23.2b (4.7)	19.6a (4.0)	18.0a (5.3)
<u>0</u> (Si)	54.1a (6.5)	53.5a (8.0)	51.3a (7.8)	52.6a (7.5)	26.1b (8.2)	39.4c (10.8)
<u>DEP</u>	26.4a (3.9)	27.3a (5.3)	22.5b (6.7)	28.3a (4.2)	7.7c (6.2)	18.6b (8.8)

Note. All scales are reported as raw scores. Mean scores with common subscripts are not significantly different (alpha = .05) using the Scheffe procedure. Standard deviations are enclosed in parentheses. Dep + Val Info = fakers receiving information about symptoms of depression and information about MMPI-2 validity indices; Dep Info = fakers receiving information about symptoms of depression; No Info= fakers who were not given symptom or validity scale information. a_n = 66. b_n = 43. c_n = 41. d_n = 49. e_n = 50. f_n = 40.

Table 3.6

Interaction of Presence or Absence of Validity Information with Presence or Absence of Depression Information for each MMPI-2 Validity Scale and Index by Participant Gender.

Scale	Men	Women
	<u>F</u> (1,167)	<u>F</u> (1,195)
<u>F</u>	1.78	7.17**
<u>Fb</u>	5.50*	7.24**
<u>F(p)</u>	3.80	7.83**
<u>Ds-2r</u>	.17	11.38***
<u>Ds-2</u>	.32	9.09**
<u>O-S</u>	.08	8.33**
<u>F-K</u>	.35	8.69**

Note. Analyses were performed using raw scores for all scales and indices.

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

Table 3.7

ANOVA Results for Comparisons of the Six Groups for each MMPI-2 Validity Scale and Index by Participant Gender.

	Men	Women
Scale	<u>F</u> (5,252)	<u>F</u> (5,283)
<u>F</u>	38.57*	50.88*
<u>Fb</u>	40.16*	59.64*
<u>F(p)</u>	22.46*	31.30*
<u>Ds-2r</u>	44.29*	55.29*
<u>Ds-2</u>	44.99*	57.11*
<u>O-S</u>	37.41*	53.18*
<u>F-K</u>	39.11*	46.92*

Note. Analyses were performed using raw scores for all scales and indices.

* $p < .0000001$

Table 3.8

Validity Scales and Indices: Means, Standard Deviations, and Differences Between Groups (Male Participants Only).

Scale	Faking Groups			Non-Faking Groups		
	Dep+Val ^a	Dep ^b	Val ^c	No Infod	Honest ^e	Clinical ^f
<u>F</u>	19.9 ^{ab} (8.0)	25.4 ^{bc} (9.7)	19.0 ^a (9.7)	28.4 ^c (10.0)	6.5 ^d (5.3)	12.0 ^d (7.8)
<u>F^b</u>	16.4 ^a (7.1)	18.7 ^{ab} (6.7)	16.1 ^a (8.1)	23.6 ^b (6.5)	4.6 ^c (4.9)	10.5 ^d (7.0)
<u>F-K</u>	8.7 ^a (10.3)	16.9 ^{bc} (11.5)	9.3 ^{ab} (11.7)	19.5 ^c (10.1)	-6.6 ^d (8.1)	-1.5 ^d (11.0)
<u>F(p)</u>	5.2 ^a (3.2)	6.7 ^{ab} (4.5)	5.2 ^a (3.6)	9.1 ^b (5.1)	2.2 ^c (2.2)	2.2 ^c (2.6)
<u>O-S</u>	158.3 ^a (65.1)	216.0 ^{bc} (61.7)	172.0 ^{ab} (68.2)	224.4 ^c (48.4)	58.0 ^d (64.1)	101.2 ^d (96.0)
<u>Ds-2</u>	28.2 ^a (8.7)	33.1 ^{ab} (8.2)	29.9 ^a (8.7)	36.3 ^b (8.0)	14.6 ^c (6.0)	18.9 ^c (9.0)

Table 3.8

Continued.

Scale	Faking Groups			Non-Faking Groups		
	Dep+Val ^a	Dep ^b	Val ^c	No Infod	Honeste	Clinicalf
<u>Ds-r2</u>	16.7 ^a (5.1)	19.6 ^{ab} (4.8)	18.2 ^{ab} (5.1)	21.6 ^b (4.4)	8.6 ^c (4.3)	10.9 ^c (6.1)

Note. All scales are reported as raw scores except for the O-S index which is a difference between T-scores. Mean scores with common subscripts are not significantly different ($\alpha = .05$) using the Scheffe procedure. Standard deviations are enclosed in parentheses. Dep + Val Info = fakers receiving information about symptoms of depression and information about MMPI-2 validity indices; Dep Info = fakers receiving information about symptoms of depression; Val Info = fakers receiving information about MMPI-2 validity indices; No Info = fakers who were not given symptom or validity scale information. F = Infrequency scale; Fb = Back F scale; F-K = F-K index; F(p) = Infrequency scale for psychopathology; O-S = Obvious-Subtle Index; Ds-2 = Dissimulation; Ds-r2 = 34 items on the Ds-r that were retained on the MMPI-2.

^an = 46. ^bn = 37. ^cn = 48. ^dn = 40. ^en = 45. ^fn = 42.

Table 3.9

Validity Scales and Indices: Means, Standard Deviations, and Differences Between Groups (Female Participants Only).

Scale	Faking Groups			Non-Faking Groups		
	Dep+Val ^a	Dep ^b	Val ^c	No Infod	Honest ^e	Clinical ^f
<u>F</u>	19.9 _a (8.0)	25.3 _a (11.4)	19.0 _a (8.7)	32.0 _b (11.5)	6.1 _c (4.1)	12.2 _d (8.2)
<u>Fb</u>	18.9 _{ab} (6.6)	22.3 _{ac} (7.8)	16.2 _{be} (8.2)	25.3 _c (7.8)	3.8 _d (3.8)	11.1 _e (8.0)
<u>F-K</u>	9.4 _a (9.3)	14.0 _a (13.6)	9.2 _a (10.4)	23.3 _b (11.7)	-7.3 _c (8.3)	-.02 _c (12.1)
<u>F(p)</u>	5.4 _a (3.6)	7.3 _a (4.7)	5.0 _a (3.0)	10.2 _b (5.1)	1.9 _c (1.8)	2.9 _c (3.0)
<u>O-S</u>	166.0 _a (54.1)	177.8 _a (66.6)	171.5 _a (64.1)	231.8 _b (49.9)	37.3 _c (64.3)	99.4 _d (93.7)
<u>Ds-2</u>	32.3 _a (7.6)	34.2 _a (9.7)	31.7 _a (10.4)	41.3 _b (8.3)	15.0 _c (7.4)	20.6 _c (10.0)

Table 3.9

Continued.

Scale	Faking Groups			Non-Faking Groups		
	Dep+Vala	Depb	Valc	No Infod	Honeste	Clinicalf
<u>Ds-r2</u>	20.3a (4.6)	20.6a (5.9)	19.4a (6.6)	25.0b (4.8)	9.2c (4.7)	12.4c (6.6)

Note. All scales are reported as raw scores except for the O-S index which is a difference between T-scores. Mean scores with common subscripts are not significantly different ($\alpha = .05$) using the Scheffe procedure. Standard deviations are enclosed in parentheses. Dep + Val Info = fakers receiving information about symptoms of depression and information about MMPI-2 validity indices; Dep Info = fakers receiving information about symptoms of depression; Val Info = fakers receiving information about MMPI-2 validity indices; No Info = fakers who were not given symptom or validity scale information. F = Infrequency scale; Fb = Back F scale; F-K = F-K index; F(p) = Infrequency scale for psychopathology; O-S = Obvious -Subtle Index; Ds-2 = Dissimulation; Ds-r2 = 34 items on the Ds-r that were retained on the MMPI-2.

a_n = 46. b_n = 37. c_n = 48. d_n = 40. e_n = 45. f_n = 42.

Table 3.10

Optimal Cutting Scores, Percent Correctly Classified, and Average Hit Rates for MMPI-2 Validity Indices for Male Patients and Male Fakers in the Three Best Faking Groups.

Index	Optimal Cutting Score	% Correctly Classified		
		Fakers (n=131)	Patients (n=42)	Average Hit Rate
<u>F</u>	14	76	67	72
<u>Fb</u>	14	71	76	74
<u>F(p)</u>	3	77	71	74
<u>F-K</u>	3	74	71	73
<u>Ds-2</u>	22	82	71	77
<u>Ds-r2</u>	14	82	79	81
<u>O-S</u>	98	88	57	72

Note. All validity scale and index cutting scores are reported as raw scores. Best fakers = those who received information about MMPI-2 validity indices, those who received information about both symptoms of depression and MMPI-2 validity indices, and those who received information about symptoms of depression. F = Infrequency scale; Fb = Back F scale; F-K = F-K index; F(p) = Infrequency scale for psychopathology; O-S = Obvious -Subtle Index; Ds-2 = Dissimulation; Ds-r2 = 34 items on the Ds-r that were retained on the MMPI-2.

Table 3.11

Optimal Cutting Scores, Percent Correctly Classified, and Average Hit Rates for MMPI-2 Validity Indices for Female Patients and Female Fakers in the Three Best Faking Groups.

Index	Optimal Cutting Score	% Correctly Classified		
		Fakers (n=150)	Patients (n=40)	Average Hit Rate
<u>F</u>	14	83	65	74
<u>Fb</u>	13	79	63	71
<u>F(p)</u>	4	68	68	68
<u>F-K</u>	2	83	63	73
<u>Ds-2</u>	20	93	53	73
<u>Ds-r2</u>	14	90	58	74
<u>O-S</u>	100	85	55	70

Note. All validity scale and index cutting scores are reported as raw scores. Best fakers = those who received information about MMPI-2 validity indices, those who received information about both symptoms of depression and MMPI-2 validity indices, and those who received information about symptoms of depression. F = Infrequency scale; Fb = Back F scale; F-K = F-K index; F(p) = Infrequency scale for psychopathology; O-S = Obvious-Subtle Index; Ds-2 = Dissimulation; Ds-r2 = 34 items on the Ds-r that were retained on the MMPI-2.

Table 3.12

Percentage of Participants Correctly Classified using Cutting Scores for MMPI-2 Validity Indices that Represent the 90th Percentile in a Large Inpatient Sample.

Scales	Validity Info		Val + Dep Info		Dep Info		No Info	
	Male (n=48)	Female (n=41)	Male (n=46)	Female (n=66)	Male (n=37)	Female (n=43)	Male (n=40)	Female (n=49)
<u>F</u> > 23	33	24	30	21	60	51	60	80
<u>Fb</u> > 19	35	34	41	46	43	70	75	76
<u>F(p)</u> > 5	38	49	44	39	54	54	75	82
<u>F-K</u> > 15	27	20	26	20	54	40	63	76
<u>Ds-2</u> > 34	33	42	30	44	54	58	63	80
<u>O-S</u> > 237	15	20	11	8	35	20	40	53

Note. All validity scale and index cutting scores are reported as raw scores. Dep + Val Info = fakers receiving information about symptoms of depression and information about MMPI-2 validity indices; Dep Info = fakers receiving information about symptoms of depression; Val Info = fakers receiving information about MMPI-2 validity indices; No Info = fakers who were not given symptom or validity scale information. F = Infrequency scale; Fb = Back F scale; F-K = F-K index; F(p) = Infrequency scale for psychopathology; O-S = Obvious -Subtle Index; Ds-2 = Dissimulation Scale.

Table 3.13

Percentage of Fakers Who Successfully Malingered by Gender and Group.

Group	Method One		Method Two	
	Men	Women	Men	Women
Dep + Val	42	34	43	38
Val Info	39	24	50	29
Dep Info	35	19	35	26
No Info	5	8	5	10

Note. Method one criteria for "successful malingering": elevation of both scale 2 and the DEP Content Scale ≥ 65 , and no elevations on any MMPI-2 validity index using cutting scores that represent the 90th percentile in a large inpatient sample. Method two criteria for "successful malingering: elevation of ≥ 65 on scale 2, 4, 6, 7, or 8, and no elevations on any MMPI-2 validity index using cutting scores that represent the 90th percentile in a large inpatient sample. Dep + Val Info = fakers receiving information about symptoms of depression and information about MMPI-2 validity indices; Dep Info = fakers receiving information about symptoms of depression; Val Info = fakers receiving information about MMPI-2 validity indices; No Info = fakers who were not given symptom or validity scale information.

Table 3.14

Application of Criteria for Successful Faking to the Clinical Group: Percentage of Patients Correctly Identified as Non-Malingers.

Group	Method One		Method Two	
	Men	Women	Men	Women
Clinical	52	28	71	63

Note. Method one criteria for "successful malingering": elevation of both scale 2 and the DEP Content Scale ≥ 65 , and no elevations on any MMPI-2 validity index using cutting scores that represent the 90th percentile in a large inpatient sample. Method two criteria for "successful malingering: elevation of ≥ 65 on scale 2, 4, 6, 7, or 8, and no elevations on any MMPI-2 validity index using cutting scores that represent the 90th percentile in a large inpatient sample. Clinical = depressed inpatient group.

CHAPTER IV

DISCUSSION

The discussion will be presented in the following order: (a) purpose of the current study, (b) summary of results, (c) comparisons to previous research, (d) issues raised by the current study, (e) limitations, and (f) suggestions for future MMPI-2 malingering research.

Purpose of study

The purpose of this study was to examine the effect of providing depression symptom information and validity scale information on malingering performance on the MMPI-2. Participants provided with this information served as an analog to the most sophisticated and motivated malingerers in clinical and forensic settings (i.e., individuals who prepare to fake by learning about the symptoms of a mental disorder or about the MMPI-2 itself). In other words, by providing information about symptoms of depression and about how the MMPI-2 detects malingering, the faking participants in this study were believed to be as well-prepared to fake depression as the most motivated, intelligent, and sophisticated malingerer. The use of well-prepared and informed participants who could probably fake depression as well as or better than most of their real-life counterparts enabled this study to be a stringent test of the MMPI-2's ability to detect the malingering of depression.

Summary of Results

Group differences on the clinical scales, the DEP content scale, and the validity indices were the primary focus of the statistical analyses.

Clinical scales and the DEP content scale

On the clinical scales, it was predicted that the group receiving validity scale information would obtain lower overall clinical scale elevations in comparison to the three other faking groups. This prediction was inaccurate because, in many cases, the other faking groups performed similarly to the validity information group. Male faking groups were more successful in obtaining clinical scale and DEP content scale elevations that were similar to those of the group of inpatients diagnosed with depression than the female fakers were (for men, on 7 scales; for women, on 5 scales). In general, the faking groups provided with information achieved lower elevations than the groups receiving no information, although, in many cases, the differences in elevations were not statistically significant. Although there was a trend showing that groups receiving validity information, whether in isolation or in combination with depression information obtained lower overall clinical scale elevations, other faking groups often did not differ significantly from those two groups. Most fakers were able to elevate both scale 2 and the DEP content scale to a T-score of 65 or above.

Validity indices

On most of the validity indices, the faking groups obtained significantly higher elevations than the inpatient comparison group. The groups receiving no information consistently obtained the highest scores for validity indices. Among women fakers, the no information group consistently obtained significantly higher elevations than the three other faking groups while the male fakers in this group consistently obtained the highest elevations, although their scores were not significantly different from some of the other faking groups. These results suggest that information about either depression or the validity scales

was helpful in allowing faking participants to achieve lower validity scale elevations than the faking groups without information.

Even though most of the faking groups in this study achieved significantly higher group means on the validity indices than the clinical comparison group, the question remains as to the clinical significance of the differences. More specifically, if an individual scores 20 T-score points higher on the F scale than the mean obtained by inpatients (as did some of the fakers in the current study), would that person routinely be judged as malingering or is that person within the expected range of scores obtained by inpatients? Group means tell us nothing about the success of an individual person.

More clinically relevant comparisons involve looking at the participants on an individual basis within groups because in a clinical setting, decisions about malingering must be made on an individual basis, not on a group basis.

Overall malingering success rates

The most important question to be answered from this research is how many participants were able to successfully malingering depression. This is a difficult question to answer because there is no established standard or objective criterion for making such a decision. Furthermore, in clinical and forensic settings, decisions about malingering are typically made using clinical interview data and other instruments along with the MMPI-2.

Several criteria were chosen in assessing malingering success in this study. The first method required that successful malingerers elevate both scale 2 and the DEP content scale over a T-score of 65 and that they not elevate any single validity index above the cutting scores that correctly classified 90% of a large inpatient sample (see Greene, 1997). According to these criteria, a large number

of fakers were judged to be successful malingerers of depression. The highest rates of success among the male participants was 42% (Validity information group) and among the female participants was 34% (Validity and Depression information group). The groups receiving no information had low rates of malingering success (for men, 5%; for women, 8%). When applying these criteria to the clinical comparison group (individuals with diagnoses of depression), a startlingly low number were correctly classified (for men, 52%; for women, 27.5%). In examining individual profiles, it became clear that depressed persons do not always elevate scale 2 and the DEP content scale (at least not in this sample). A high number of female inpatients did not elevate scale 2 or the DEP content scale, thus accounting for the low percentage of correct classification.

Because my method of identifying successful fakers misclassified so many of the bona fide patients, the criteria I used for judging successful malingering performance were judged to be inadequate. Individuals in the clinical sample had been diagnosed as having depression but did not obtain elevations on the two scales for which one would intuitively expect elevations. An examination of group means for the clinical sample revealed elevations (T-score > 65) for all clinical scales except scale 9 for men and scales 1, 3, 9, and 0 for women. An examination of MMPI-2 codetypes of the individual inpatients revealed a heterogeneous picture. Although the traditional MMPI-2 codetypes typical of depressed individuals were present, there were many other codetypes represented in the sample (i.e., 6/8-8/6). (See Table 4.1. at the end of the Discussion section for a list of the two most frequent codetypes for each group by gender.)

A less stringent method of determining successful malingering was employed that used the same validity index cutting scores, but used different, more lax criteria for

clinical scale elevations that reflected better the results of the inpatient sample. Specifically, an elevation of scale 2, 4, 6, 7, or 8 (the highest mean elevations achieved by the inpatients) was used as the clinical scales' criterion for successful malingering. Using this method, more fakers were able to elude detection (for men, 50% of the validity information group; for women, 38% of the validity and depression information group), and more of the inpatient group was correctly classified (for men, 71%; for women, 62.5%). Even using the lax criteria, an unacceptably high number of bona fide depressed inpatients were not correctly identified as exhibiting depression on the MMPI-2.

On the MMPI-2, detection of malingering is primarily the job of the validity indices. Indeed, most of the fakers who were deemed "unsuccessful malingerers" were detected by the validity scales and not by the clinical scales' criterion. It is extremely difficult, therefore, to choose a standard for the clinical scales to aid in detecting malingering that would be accurately reflective of the scores obtained by bona fide depressed inpatients and that would contribute much in detecting malingering beyond the validity indices. Using this clinical sample as an example, almost any MMPI-2 clinical scales profile pattern can be obtained by a genuinely depressed inpatient. Therefore, it is almost impossible to predict a profile pattern of a depressed person on the individual level with any accuracy. In reality, in most cases, a genuinely depressed person's diagnosis of depression would not be based solely on MMPI-2 results unless the validity indices suggested malingering (or unless the MMPI-2 results were pointing to the likelihood of psychosis).

Due to the inherent problems in defining criteria for successful malingering, the percentages of groups who were reported to be successfully malingering depression in the previous paragraphs should be considered tentative.

Comparisons to Previous Research

Detecting malingering or judging the successfulness of fakers can be approached from different perspectives. For example, many MMPI-2 malingering studies have inferred successful malingering by comparing group means (fakers vs. patients, or fakers vs. honest responders). Other studies have established optimal cutting scores that most efficiently differentiated groups or applied cut-off scores that have been suggested for use in MMPI/MMPI-2 textbooks.

Cutting scores

Due to the wide variety of methods used to determine successful malingering on the MMPI-2, it is difficult to compare studies, particularly when cutting scores are established that apply only to the particular sample used in the study. For example, in Rogers et al. (1993), a study using a similar design to the present experiment (but asking individuals to fake schizophrenia), the authors established cutting scores that most efficiently differentiated honest responders (controls and schizophrenics) from different categories of community simulators. Based on these cutting scores, it was reported that scales Ds-r2 and O-S were superior to the other validity indices in terms of correctly classifying more participants who were given information about the validity indices on the MMPI-2 (the most successful faking group in this study) and a large percentage of other faking groups.

In the current study, optimal cutting scores were reported which were quite different than the cutting scores which were most efficient in the Rogers et al. (1991) study. This difference in cutting scores was most likely due to the fact that different patterns of psychopathology were being studied (schizophrenia vs. depression). In the current study, for example, a raw cutting score of greater than 13 for the F scale differentiated the best faking groups from

the clinical comparison group for both men and women. This cutting score accurately identified 76% of male fakers, 83% of female fakers, 67% of male inpatients, and 65% of female inpatients. In the Rogers et al. (1993) study, an F scale cutting score of greater than 28 most efficiently differentiated between groups (97% of schizophrenic patients correctly classified, 47% of fakers receiving information correctly classified). In the current study, the classification accuracy achieved for the clinical group using optimal cutting scores was typically much lower (ranged from 53% to 79%) than that seen in Rogers' study for schizophrenic inpatients (ranged from 81% to 97%).

There is a large range of cutting scores reported across MMPI-2 malingering studies. This variation is probably caused by a number of factors including type of psychopathology being faked and different faking instructions. For example, prison populations and psychotic populations are known to elevate the F scale more than other groups (Berry et al., 1991). Underlying differences in expected F scores would clearly influence optimal cutting scores for the F scale. Berry et al. (1991) suggested that, rather than searching for the best cutting score to detect malingering for each index, that identifying cutting scores for malingering that apply to different populations (disability applicants, prisoners, inpatients, etc.) may be a more productive avenue of research.

Another reason for the wide range of cutting scores for MMPI-2 validity indices is the substantial overlap between the distributions of scores for the two criterion groups. When there is substantial overlap of scores among the groups, then the optimal cutting score can vary greatly depending on the specific samples used in the study (Rogers et al., 1994).

Cutting scores for the F scale have ranged from 17 to 30 in previous malingering studies (Rogers et al., 1994).

Rogers et al. (1993) pointed out that individual clinicians must decide what kinds of classification errors are acceptable (e.g., malingerers classified as patients or patients classified as malingerers). Higher cutting scores that minimize the misclassification of genuine patients probably make more sense in most clinical settings given the extremely negative effect of a judgment that malingering has occurred (Rogers et al., 1993).

Group means

Comparison to Bagby et al. (1997)

Some of the results of this study are similar to the results found in a recent study (Bagby et al., 1997) when similar groups from each study are compared. In that study, participants ($n = 40$; 7 men and 33 women) were asked to fake schizophrenia or to fake depression on the MMPI-2. Information about the disorders from the DSM-IV and financial incentives for successful faking were provided to participants. The following comparison will include only the faking groups in the current study that received depression symptom information since these are the groups that are the most similar to the group faking depression in the Bagby et al. (1997) study. In the current study, the female participant group receiving depression information achieved almost identical scale 2 elevations ($M = 93.3$) as the depression group in the Bagby et al. (1997) study ($M = 94.0$). Although the number of women in the Bagby et al. (1997) depression group was not reported, it can be surmised that it was approximately 75% women. In terms of raw score elevations on the validity indices, the female participants in the depression information group in the current study scored lower on every index (F, Fb, Fp, F-K, O-S, and Ds-2) in comparison to the depression group in Bagby's study (e.g., on the Fb scale, raw scores of 22.3 and 24.6 respectively; on the O-S index; scores of 177.8 and 213.9

respectively). The male participants in the depression information group in the current study also scored lower than the faking depression group in the Bagby et al. (1997) study with the exception of the O-S index and the F-K index on which both groups scored similarly.

There are three possible explanations for the lower scores for validity indices among the participants in the current study compared to the scores of participants faking depression in the Bagby et al. (1997) study, and all three explanations are associated with the different methods in the two studies. In the current study, participants in the depression symptom information group received an easy-to-understand summary of the DSM-IV diagnostic criteria for depression, whereas in the Bagby et al. (1997) study, participants were given the diagnostic criteria for major depression directly from the DSM-IV. Furthermore, in the current study, participants were forced to process the information through worksheets and quizzes, whereas in the Bagby et al. (1997) study, there was no mention of such a procedure. Due to these differences in procedure, the participants in the current study may have thoroughly processed the fact that an individual does not have to have all of the DSM-IV symptoms for depression in order to receive a diagnosis of depression. Items on the worksheet and the quiz referred to this idea. Endorsing fewer symptoms of depression would lead to lower scores on validity indices because as individuals report fewer symptoms of distress, the scores on validity indices such as F tend to decrease.

The second difference in method that probably contributed to the differences in mean validity indices scores across the two studies involves the instructions given to participants. In the Bagby et al. (1997) study, students were given "brief instructions for the study that gave examples of different situations under which feigning

might occur" (p. 652) whereas in the current study, a detailed college-specific scenario was given to participants that stressed the point that the depressive symptoms should not be overexaggerated. Clearly, this could account for participants' lower scores on the validity scales in the current study in comparison to the fakers of depression in the Bagby et al. (1997) study.

The third difference in method that may have contributed to the differences in mean scores for validity indices across the two studies involves the level of motivation of the participants. The scenario in the present study was designed to increase the motivation of participants by depicting a situation that many college students could relate to (financial pressures, failing grades, scholarship, dealing with parents). Presumably, if the participants identified with the scenario, they might be expected to become more involved in the experiment and more motivated to successfully fake depression on the MMPI-2. Apparently, Bagby et al. (1997) did not offer a scenario that was involving, but rather, provided several examples of "different situations under which feigning might occur" (p. 652).

Comparison to Rogers et al. (1993)

Because the study by Rogers et al. (1993) is one of the only MMPI-2 malingering studies that has a similar design to the current study, results will be compared between these two studies in some detail. In terms of raw scores for the validity indices, the most striking difference across the two studies is that the group receiving validity scale information in isolation in the Rogers et al. (1993) study scored much lower on all validity scales (F, Fb, F-K, O-S, and Ds-2r) than their male and female counterpart groups in the current study. For example, the validity information group in Rogers' study obtained a score of 118.2 on the O-S

index, whereas the men and women in the validity information group in the present study obtained O-S scores of 172.0 and 171.5 respectively. The reverse pattern was seen on three validity indices in comparing the groups receiving both symptom and validity information in the two studies. This group scored much higher on F, F-K, and O-S in Rogers' study than both the men and women in the analogous group in the current study. In the Rogers et al. (1993) study, for some reason, the group receiving both schizophrenia and validity information scored significantly higher on the validity indices than the group receiving only validity information. In the current study, the groups receiving both depression information and validity scale information scored often similarly on the validity indices to the groups receiving only validity scale information. Rogers et al. (1993) guessed that the symptom information served to distract participants from the validity information, resulting in higher scores. This phenomenon was not observed in the current study. Perhaps the severity of schizophrenia symptoms in comparison to depression symptoms was more distracting for participants in the Rogers' et al. (1993) study than for the participants in the current study. However, there is no clear explanation for the differences in validity indices elevations between the groups receiving validity information and symptom information in the two studies.

The differences in results when comparing groups receiving symptom information in the two studies may be a result of the kinds of disorders that participants were asked to fake. Faking schizophrenia would naturally result in higher elevations on the F scale, for instance, since many items related to bizarre or psychotic experiences load onto the F scale.

Comparison to Lamb et al. (1994)

Lamb et al. (1994) also examined the effects of providing symptom information (on closed head injury) and MMPI-2 validity information on malingering success on the MMPI-2. In the Lamb et al. (1994) study, mean elevations on two validity scales, F and Fb, were reported by group (men and women combined). All faking participant groups in the Lamb et al. (1994) study scored much lower on the two validity scales than their counterpart groups, both male and female, in the current study. For example, the mean F scale T-score was 65.6 for the validity information group in the Lamb et al. (1994) study, whereas it was much higher for the male and female groups in the current study (for men, M = 91.0; for women, M = 97.2). In the Lamb et al. (1994) study, the highest scoring faking group on F and Fb was the closed head injury information group (for F, M = 89.5; for Fb, M = 90.8). In contrast, the highest scoring faking group in the current study was the no information group (for men, F scale mean = 109.9 and Fb scale mean = 116.7; for women, F scale mean = 114.7 and Fb scale mean = 115.0). Although the mean scores on the two validity scales are much lower across the board for the Lamb et al. (1994) faking groups, a similar pattern of scores among the faking groups receiving information is observed in the two studies. However, when the no information groups in the two studies are compared to each other and to the groups receiving information, a different pattern emerges. In the current study, the no information groups scored much higher than the depression information group, whereas in the Lamb et al. (1994) study, the closed head information group scored much higher than the no information group.

The difference in findings regarding the performance of the no information group in the two studies most likely relates to the nature of the disorder that the participants were being asked to fake. Whereas most people are

reasonably well-informed about the symptoms of depression and may have even experienced some of those symptoms, the same cannot be said of an individual's knowledge of the common effects of closed head injury on cognitive abilities and on personality. By nature of the design used in the two studies, the no information group was not given any type of information about the disorder they were being asked to fake. In the present study, the no information group probably had a general idea about the symptoms of depression, but in the Lamb et al. (1994) study, it is unlikely that the no information group even had a general idea about how to portray a person with closed head injury. Therefore, in the Lamb et al. (1994) study, the lack of information about the common impairments associated with closed head injury may have lowered much of the no information faking group's scores because participants were unsure of what MMPI-2 items to endorse. In the current study, however, the lack of symptom information about depression may have had the opposite effect. Because it is likely that participants in the no information group had an idea of what depression was, even if erroneous, these ideas may have led participants in this study to have overreported symptoms in relation to the participants in the no information group in the Lamb et al. (1994) study who probably had very little idea of what symptoms constitute a disorder (e.g., closed head injury). Endorsement of more symptoms of distress on the MMPI-2 leads to higher scores on the validity scales and indices and vice versa. In conclusion, it can be said that in the current study symptom information served to lower validity and clinical scale elevations in comparison to the group receiving no information, whereas in the Lamb et al. (1994) study, the opposite effect was seen.

Interestingly, the groups in the Lamb et al. (1994) study and in the current study that received both kinds of

information were able to lower their mean score on F and Fb to a level only slightly above the groups receiving validity scale information only. This is contrary to the results of Rogers et al. 1993) in his study of malingering schizophrenia on the MMPI-2 in which he found that the group receiving both kinds of information scored much higher on the validity indices in comparison to the group receiving only validity information (see Rogers' section above).

Looking at the overall results from the current study and the Lamb et al. (1994) study shows that the effect of providing symptom information on faking participants' performance on the MMPI-2 (relative to other faking groups' performance) is dependent on the type of disorder that participants are asked to fake. In the case of closed head injury, providing symptom information without validity scale information caused participants to overendorse symptoms in comparison to all the other faking groups in the study (most notably, achieving much higher scores than the fakers who received no information). In the current study, providing symptom information without validity scale information caused participants to endorse fewer symptoms than naive fakers. These results suggest that providing depression information to persons asked to fake depression is helpful in their successful malingering of depression on the MMPI-2, whereas providing closed head injury information to persons asked to fake that condition causes more people to be detected as a malingerer by the MMPI-2 validity scales. Apparently, providing information about the disorder that the person is being asked to fake on the MMPI-2 has a differential effect on their malingering performance based on the nature of the specific disorder that participants are being asked to fake. A limitation of the Lamb et al. (1994) study was that it did not use a clinical comparison group, instead comparing the faking groups to a group responding honestly to the MMPI-2. Without a clinical comparison

group, cutting scores for the validity indices that distinguish between fakers and bona fide patients cannot be established. In addition, it is more difficult to judge successful malingering of a disorder if a clinical comparison group is not used.

Issues Raised by the Current Study

Nature of depression information

Participants in the current study were forced to process the depression information thoroughly. One of the key points emphasized in the depression information, worksheet, and quiz was that a person did not have to exhibit all the symptoms of depression to receive a depression diagnosis. One of the tips contained on the validity information sheet suggested that participants pick a few symptoms and "stick with them." The point made with the depression symptom information that all symptoms did not have to be endorsed to receive a major depression diagnosis may have made the validity information and the depression information not totally distinct from each other. However, to represent the complete criteria for major depression from DSM-IV, the information about number of symptoms necessary for diagnosis had to be included. Therefore, there was no other way to present this information without creating a slight overlap with the validity information. There is a possibility that the depression information as it was presented may have inclined the group receiving depression information only to be aware that they should not endorse too many items to successfully fake depression on the MMPI-2. The overlap of depression information and validity scale information may have led to the differences in elevations for the symptom information group in the current study in comparison to other similar studies (Lamb et al., 1994; Rogers et al., 1993). For example, in the Lamb et al. (1994) study, the symptom information groups scored much

higher than the other information groups on validity indices, whereas in the current study, the depression information group sometimes achieved similar elevations on the validity indices as the other information groups.

Two feedback questionnaire items provide a note of caution in assuming that the overlap of depression information and validity scale information in the current study led to a difference between the scores for the symptom information groups and the other faking groups in the current study that was smaller than the equivalent differences in previous studies. Two self-report feedback questionnaire items assessing the strategies that participants used for malingering indicated that the depression information groups were no more likely to limit depression symptom endorsement than any other group, including the no information group.

Motivation of fakers

The purpose of the current study was to create an analog for the most sophisticated, intelligent, and motivated malingerers. There is no doubt that some individuals prepare to malingering either by attempting to study symptoms of the disorder they are wanting to fake or exaggerate, or by studying the MMPI-2 itself. In addition, the literature indicated that some unscrupulous lawyers have actually coached their clients on strategies to successfully malingering on the MMPI-2. A person wanting to malingering a mental illness on the MMPI-2 is likely to have an important outcome at stake whether it be a cash settlement in a forensic case, disability payments in a disability applicant case, or the right to reside in a psychiatric prison in a case involving a criminal. Clearly, persons who decide to malingering are very motivated to succeed because the potential rewards are very large.

One problem with MMPI-2 malingering research is that it is very difficult to create a situation that comes close to simulating the motivation levels that are present in a real-life malingerer. To deal with this problem, my study used financial incentives, an involving scenario, and strict exclusion rules so that unmotivated participants would not be included in the results. The summary of the strategies and the symptoms of depression are the kinds of information that a motivated and sophisticated malingerer would seek and would be able to obtain.

In this study, motivation was assessed using several methods. One method was a question on the feedback questionnaire that asked, "How motivated were you to try to appear depressed on the MMPI-2?" In the Likert response format, participants were asked to circle a number between one (did not try at all) and five (very motivated). The mean score on this question across participants in the faking groups who were included in the data analysis was 4.2, indicating a high level of motivation. Participants who were dropped from the analyses due to low quiz performance, high VRIN scores, or leaving too many items blank on the answer sheet were significantly less motivated than the persons who were not excluded according to the feedback questionnaire item assessing motivation (for excluded, $\bar{M} = 3.5$).

Although it is almost impossible to create a situation that closely approximates the sort of incentive that a real-life malingerer has, this experiment was quite successful in motivating participants as shown in the motivation feedback questionnaire item. Because it was expected that some participants would not be motivated to perform at their best, methods were built into the experiment to "weed out" such participants so that the findings would not be contaminated or diluted. Because the faking groups receiving information typically performed differently than

the faking groups receiving no information, it can also be concluded that participants attended to the information supplied and used that information to guide their responses on the MMPI-2. Differential performance among the faking groups is further evidence that the participants were invested in the experiment. Additionally, significantly higher mean quiz scores by the groups receiving information in comparison to the honest groups (who did not receive information to prepare them to take the quizzes) indicate that faking participants receiving information appear to have been motivated enough to attend to the information provided as demonstrated by their quiz scores.

Gender differences

The main focus of the current study was not on differences in the MMPI-2 scores of men and women. Data was analyzed separately because men and women are known to respond differently to measures of depression. These differences are reflected in the gender differences seen on some of the MMPI-2 scales and indices. Although examining gender differences was not a goal of this study, one gender difference bears mentioning. Overall, the male fakers were more successful than the female fakers. This is a puzzling finding and one lacking a ready explanation. There were big differences in the male and female inpatient groups as well, with the women inpatients displaying significantly fewer elevations on scale 2 and the DEP content scale in relation to the male inpatients. The differences between men and women in the inpatient group are also difficult to understand and may have been specific to the particular sample used in this study.

Limitations

Base rates

This study has several important limitations. In establishing the optimal cutting scores and using discriminant analyses, the issue of the base rate of malingering was not addressed.

Base rates of disorders are important to take into consideration because the probability of classifying an individual as having a disorder based on a positive test score is dependent on the base rate of the target disorder in the population that is being evaluated (Elwood, 1993). When assessing for disorders that are rare (i.e., have low base rates), it is difficult even with a highly valid measure to classify an individual as having the target disorder with an accuracy any greater than chance. For example, if a test has a 90% accuracy rate of identifying a disorder, but the disorder occurs in only 5% of the population being assessed, it would identify nearly all (90%) of the individuals who actually have the disorder, but most (68%) of the individuals identified by the test as having the disorder would not (false positives).

The base rate issue is ignored in most MMPI-2 malingering research. Part of the reason that base rates of malingering are not considered is because there are no established and reliable estimates of the base rate of malingering. Due to the very nature of "malingering," it is difficult to estimate the prevalence. In addition, there is probably much variability in the base rate of malingering in different settings.

In the current study, the overall rates of correct classification of the best faking groups and the clinical group based on the optimal cutting scores were misleading because the base rate of malingering was not taken into consideration. When base rates for malingering are used in figuring correct classification, there will be a dramatic

decrease in positive predictive power assuming that the base rate for malingering is much lower than 50%. Negative predictive power (NPP) is the ratio of true negatives to all test negatives, or in other words, the probability that a person with a negative test results actually does not have the target disorder. Positive predictive power (PPP) is the ratio of true positives to all test positives, or the probability that a person with a positive test results actually has the target disorder. "Overall rates of correct classification are misleading when applied to discriminations of low base rate disorders" (p. 413, Elwood, 1993) such as malingering because the low PPP (true positives) is concealed by the high NPP (true negatives) (Elwood, 1993).

In this study, it was reported that an F scale score greater than 13 accurately classified 76% of male fakers and 67% of male patients. Base rate estimates were not used and therefore, in effect, the base rate for malingering used in the cutting score and discriminant analyses was treated as much greater than it really is (for example, 50% in the discriminant analysis). In the case of the F scale cutting score greater than 13 mentioned above, the lack of use of base rates led to artificially high estimates of PPP (.88) and NPP (.47) in predicting malingering. When base rate estimates are employed, the cutting scores are much less impressive due to lowered PPP. To illustrate how PPP and NPP are affected when base rates are used, PPP and NPP were recomputed using an estimated malingering base rate of 20%.⁹ The 20% base rate of malingering resulted in a PPP of .36 and an NPP of .92 indicating that there was a 36% probability that the male participants who scored above the

⁹ This figure (20%) may be an overestimation of the actual base rate of malingering in a disability setting, but one study reported a malingering rate of 19% in a SSI disability setting (Griffin et al., 1997).

F cutting score 13 were actual malingerers and 92% probability that the participants who scored below 13 were actually not malingering. Using the base rate for malingering produces a more realistic estimate of the F scale's accuracy for predicting malingering in other settings (Sivec et al., 1995).

The lower the base rate chosen, the lower the PPP. In order to increase PPP, a higher cutting score would have to be used. Only one MMPI-2 malingering study has attempted to employ base rates (Sivec et al., 1995). In that study, an F scale cutting score of 19 seemed to optimally separate faking subjects from bona fide patients. However, when a malingering base rate of 7.4 was employed, the PPP dropped dramatically. In order to predict malingering in settings with similar base rates of malingering, the authors indicated that a higher cutting score for F would need to be used (F > 23) (Sivec et al., 1995).

Motivation of participants

Another limitation of this study and all research studying malingering on the MMPI-2 involves motivation levels of faking participants in comparison to genuine malingerers. As mentioned earlier, real-life malingerers typically have very enticing incentives motivating their faking performance. It is doubtful that this experiment created the level of motivation that is comparable to that of a genuine malingerer. However, this is a drawback for this kind of research in general. This study appears to have done as well as or better than other similar studies in motivating participants. In addition, strict exclusion criteria were imposed so that participants with low motivation would not jeopardize the integrity of the experimental manipulation.

College student participants

Another limitation of the present study is the use of college students as participants. Data from college students may not be generalizable to other populations because students may be better able to make use of the "coaching" on validity scale information and depression information due to their higher levels of intelligence and sharpened study skills (Lamb et al., 1994). Although external validity is sacrificed, the advantage of using college students as participants is that internal validity is maximized and it is easier to have larger sample sizes, ensuring adequate statistical power. Clearly, the current findings should be replicated using other populations to determine if the findings are generalizable.

Clinical sample

To improve upon the clinical relevance of the study, a clinical comparison group of depressed inpatients was included in the study following Rogers' (1988) suggestion. The inclusion of a patient group allowed fakers to be compared not only to other normal individuals who are not faking, but also to a more clinically relevant group. In examining the codetypes achieved by the inpatient groups, it was noted that the most frequently obtained 2-point high codetype was 6-8/8-6 (for men, 14%; for women, 14%). A 2-point codetype including scale 2 was obtained by 23% of the male inpatients and by 30% of the female inpatients. A 6-8/8-6 codetype is typically not one of the most common codetypes obtained by depressed persons (Schipper, 1993). It is particularly surprising because none of the inpatients in the sample had a secondary diagnosis that would be suggestive of psychotic symptomatology. This clinical

sample may not have been typical of depressed persons.¹⁰ However, the range of codetypes displayed by the inpatient sample is a testament to the fact that many different codetypes can be obtained by depressed individuals.

Validity scale information

Another possible limitation of this study involves the easily comprehensible validity scale information provided participants in the current study. It is unlikely that most real-life malingerers have access to such uncomplicated MMPI-2 validity scale information unless the malingerer is coached by a fairly knowledgeable person or has an educational background in psychology. Motivated malingerers are probably more likely to obtain a psychology textbook on the MMPI-2 that provides technical information, but that would be quite difficult for the lay person to understand. There would have been two potential problems with providing such technical information to participants in the current study. First, it is likely that it would not be comprehended. Second, it would have increased the length of the experiment significantly since more study time would have been required. Although external validity is sacrificed, a conscious decision was made to provide information to participants that was easy to understand. In other words, providing understandable validity scale information provided a stringent test of the MMPI-2's ability to detect persons attempting to fake depression.

¹⁰This clinical sample was used in an earlier MMPI-2 study by Ben-Porath et al. (1991). The depressed inpatients were successfully differentiated from a schizophrenic sample of inpatients by the clinical scales (for men, scales 2 and 9; for women, scales 2 and 4) and by the content scales (for men, BIZ and DEP; for women, DEP and ANX).

Future Research

Future research needs to take into consideration how clinicians make decisions about malingering. Studies that merely report significant differences between group means are interesting, but provide no guidance to the clinician trying to make a malingering determination for an individual. In addition, estimated base rates for malingering should be used (if they become available in the future) to more accurately predict malingering beyond the sample being reported on.

This study should be replicated using a non-college population to determine if the present results generalize to other populations. In addition, more research should be done that examines the successfulness of the MMPI-2 in detecting the malingering of other common disorders such as generalized anxiety disorder or bipolar disorder.

Conclusion

Despite some of the limitations described above, this study was able to improve on the existing methodology used in MMPI-2 malingering research. One particular innovation was asking participants if they had ever experienced depression or known someone who had depression. The idea that a person's personal experience with depression might affect their ability to malingering depression was not confirmed by the data in the study.

This study was useful in demonstrating that information about depression and validity scales affects the performance of college students who are attempting to fake depression on the MMPI-2. Specifically, as many as 50% of the participants in one of the faking groups (male group receiving validity information) were able to elude detection by the MMPI-2. Thus, it is likely that some bona fide malingerers of depression who are motivated and sophisticated and who access similar information to that

provided in this study might also be able to elude detection on the MMPI-2. Of course, in a real-life clinical assessment situation, a determination of malingering would be based on many sources of information including a clinical interview. Therefore, if the MMPI-2 failed to detect a malingerer, there is a chance of detection by another method.

Table 4.1

Most frequent two point MMPI-2 profile codetype obtained by each group by gender.

Group	MMPI-2 codetypes	
	Most frequent	2nd most frequent
Clinical		
Men	6-8/8-6 (14%)	7-8/8-7 (10%), WNL* (10%)
Women	6-8/8-6 (13%)	7-8/8-7, 2-7/7-2, 8-6-4, Spike 2, Spike 8 (5% each)
Depression		
Men	6-8/8-6 (35%)	2-8/8-2 and 7-8/8-7 (16% each)
Women	2-8/8-2 (23%)	2-7/7-2 (21%)
Validity		
Men	7-8/8-7 (13%)	2-8/8-2 and 6-8/8-6 (10% each)
Women	2-8/8-2 (17%)	0-2/2-0 (9%)
Both		
Men	7-8/8-7 (28%)	2-7/7-2 (17%)
Women	2-7/7-2 (35%)	2-8/8-2 (14%)
No Info		
Men	7-8/8-7 (30%)	6-8/8-6 (25%)
Women	6-8/8-6 (39%)	2-8/8-2 (14%)

Table 4.1

Continued.

Group	MMPI-2 codetypes	
	Most frequent	2nd most frequent
Honest		
Men	WNL* (40%)	Spike 9 (9%)
Women	WNL* (44%)	Spike 9 (4%)

Note. Inpatients = inpatients who received a diagnosis of Major Depression, depression = fakers receiving information about symptoms of depression, validity = fakers receiving information about MMPI-2 validity indices, both = fakers receiving information about symptoms of depression and information about MMPI-2 validity indices, no info = fakers receiving no information, honest = participants who were instructed to answer the MMPI-2 honestly. Percentages of participants obtaining a particular codetype are listed in parentheses following the codetype.

*WNL = within normal limits profile.

REFERENCES

- American Psychiatric Association (1980). Diagnostic and statistical manual of mental disorders (3rd ed.). Washington, DC: Author.
- American Psychiatric Association (1987). Diagnostic and statistical manual of mental disorders (3rd ed., revised). Washington, DC: Author.
- American Psychiatric Association (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.
- Arbisi, P. & Ben-Porath, Y. (1995). An MMPI-2 infrequent response scale for use with psychopathological populations: The F(p) scale. Psychological Assessment, 7, 424-431.
- Austin, J. S. (1992). The detection of fake-good and fake-bad on the MMPI-2. Educational and Psychological Measurement, 52, 669-674.
- Baer, R. A., Wetter, M. W., & Berry, D. T. R. (1995). Effects of information about validity scales on underreporting of symptoms on the MMPI-2: An analogue investigation. Assessment, 2, 189-200.
- Bagby, R. M., Buis, T., & Nicholson, R. A. (1995). Relative effectiveness of the standard validity scales in detecting fake-bad and fake-good responding: Replication and extension. Psychological Assessment, 7, 84-92.
- Bagby, R. M., Rogers, R., & Buis, T. (1994). Detecting malingered and defensive responding on the MMPI-2 in a forensic inpatient sample. Journal of Personality Assessment, 62, 191-203.
- Bagby, R. M., Rogers, R., Buis, T., Nicholson, R. A., Cameron, S. L., Rector, N. A., Schuller, D.R., & Seeman, M.V. (1997). Detecting feigned depression and schizophrenia on the MMPI-2. Journal of Personality Assessment, 68, 650-664.
- Bagby, R. M., Rogers, R., Buis, T., & Kalembe (1994). Malingered and defensive response styles on the MMPI-2: An examination of validity scales. Assessment, 1, 31-38.

- Ben-Porath, Y. (1994). The ethical dilemma of coached malingering research. Psychological Assessment, 6, 14-15.
- Ben-Porath, Y., Butcher, J. N., & Graham, J. R. (1991). Contribution of the MMPI-2 content scales to the differential diagnosis of schizophrenia and major depression. Psychological Assessment, 3, 634-640.
- Berry, D. T. R., Lamb, D. G., Wetter, M. W., Baer, R. A., & Widiger, T. A. (1994). Ethical considerations in research on coached malingering. Psychological Assessment, 6, 16-17.
- Berry, D. T. R. (1995). Detecting distortion in forensic evaluations with the MMPI-2. In Y. S. Ben-Porath, J. R. Graham, G. C. N. Hall, R. D. Hirschman, & M. S. Zaragoza (Eds.), Forensic Applications of the MMPI-2 (pp.82-102). Thousand Oaks, CA: Sage Publications, Inc.
- Berry, D. T. R., Baer, R. S., & Harris, M. J. (1991). Detection of malingering on the MMPI: A meta-analysis. Clinical Psychology Review, 11, 585-598.
- Boone, D. E. (1994). Validity of the MMPI-2 depression content scale with psychiatric inpatients. Psychological Reports, 74, 159-162.
- Boone, D. E. (1995). Differential validity of the MMPI-2 subtle and obvious scales with psychiatric inpatients: Scale 2. Journal of Clinical Psychology, 51, 526-531.
- Butcher, J. N. (1990). The MMPI-2 in psychological treatment. New York: Oxford University Press.
- Butcher J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). Minnesota Multiphasic Personality Inventory-2. Manual for administration and scoring. Minneapolis: University of Minnesota Press.
- Cassisi, J. E. & Workman, D. E. (1992). The detection of malingering and deception with a short form of the MMPI-2 based on the L, F, and K scales. Journal of Clinical Psychology, 48, 54-58.
- Cohen, J. (1992). A power primer. Psychological Bulletin, 112, 155-159.
- Cunnien, A. (1988). Psychiatric and medical syndromes associated with deception. In R. Rogers (Ed.), Clinical assessment of malingering and deception (pp. 13-33). New York: Guilford Press.

- Dahlstrom, W. G., Welsh, G. S., & Dahlstrom, L. E. (1972). An MMPI handbook: Vol. I. Clinical applications (rev. ed.). Minneapolis: University of Minnesota Press.
- Dahlstrom, W. G., Welsh, G. S., & Dahlstrom, L. E. (1975). An MMPI handbook: Vol. II. Research applications. Minneapolis: University of Minnesota Press.
- Elwood, R. W. (1993). Psychological tests and clinical discriminations: Beginning to address the base rate problem. Clinical Psychology Review, 13, 409-419.
- Fox, D. D., Gerson, A., & Lees-Haley, P. R. (1995). Interrelationships of MMPI-2 validity scales in personal injury claims. Journal of Clinical Psychology, 51, 42-47.
- Graham, J. R. (1993). MMPI-2: Assessing Personality and Psychopathology (2nd ed.). New York: Oxford University Press.
- Graham, J. R., Watts, D., & Timbrook, R. E. (1991). Detecting fake-good and fake-bad MMPI-2 profiles. Journal of Personality Assessment, 57, 264-277.
- Greene, R. (1978). An empirically derived MMPI carelessness scale. Journal of Clinical Psychology, 46, 407-410.
- Greene, R. (1980). The MMPI: An interpretive manual. New York: Grune and Stratton.
- Greene, R. (1991). The MMPI-2/MMPI: An interpretive manual. Boston: Allyn and Bacon.
- Greene, R. L., Gwin, R., & Staal, M. (1997). Current status of MMPI-2 research: A methodologic overview. Journal of Personality Assessment, 68, 20-36.
- Greene, R. L. (1997). Assessment of malingering and defensiveness by multiscale inventories. In R. Rogers (Ed.), Clinical assessment of malingering and deception (2nd Ed.) (pp. 169-207). New York: Guilford Press.
- Griffin, G. A., Normington, J., May, R., & Glassmire, D. (1997). Assessing dissimulation among social security disability income applicants. Journal of Consulting and Clinical Psychology, 64, 1425-1430.
- Grow, R., McVaugh, W., & Eno, T. (1980). Faking and the MMPI. Journal of Clinical Psychology, 36, 910-917.

- Hare, R. D. (1985). Comparison of procedures for the assessment of psychopathy. Journal of Consulting and Clinical Psychology, 53, 7-16.
- Heaton, R., Smith, H., Lehman, R., & Vogt, A. (1978). Prospects for faking believable deficits on neuropsychological testing. Journal of Consulting and Clinical Psychology, 46, 892-900.
- Iverson, G. L., Frazen, M. D., & Hammond, J. A. (1995). Examination of inmates' ability to malingering on the MMPI-2. Psychological Assessment, 7, 118-121.
- Lachar, D., & Wrobel, T. A. (1979). Validating clinicians' hunches: Construction of a new MMPI critical item set. Journal of Consulting and Clinical Psychology, 47, 277-284.
- Lamb, D. (1992). Malingering closed head injury on the MMPI-2: Effect of information about symptoms and validity scales (Doctoral dissertation, University of Kentucky, 1992). Dissertation Abstracts International, 53, 1611.
- Lamb, D., Berry, D. T. R., Wetter, M., & Baer, R. (1994). Effects of two types of information on malingering of closed-head injury on the MMPI-2: An analog investigation. Psychological Assessment, 6, 8-13.
- Lechart, B. (1994). A revised dissimulation scale applicable to the MMPI-2. American Journal of Forensic Psychology, 12, 5-15.
- Lees-Haley, P. (1991a). Ego strength denial on the MMPI-2 as clue to simulation of personal injury in vocational neuropsychological and emotional distress evaluations. Perceptual and Motor Skills, 72, 815-819.
- Lees-Haley, P. (1991b). MMPI-2 F and F - K scores of personal injury malingerers in vocational neuropsychological and emotional distress claims. American Journal of Forensic Psychology, 9, 5-14.
- Lees-Haley, P. (1992). Efficacy of MMPI-2 validity scales and MCMI-II modifier scales for detecting spurious PTSD claims: F, F-K, Fake Bad scale, Ego Strength, Subtle-Obvious subscales, DIS, and DEB. Journal of Clinical Psychology, 48, 681-688.
- Lubin, B., Larsen, R., & Matarazzo, J. (1984). Patterns of psychological test usage in the United States: 1935-1982. American Psychologist, 39, 451-454.

- Lees-Haley, P. & Dunn, J. T. (1994). The ability of naive subjects to report symptoms of mild brain injury, post-traumatic stress disorder, major depression, and generalized anxiety disorder. Journal of Clinical Psychology, 50, 252-256.
- Reich, J. (1986). DSM-III diagnoses in Social Security Disability applicants referred for psychiatric evaluation. Journal of Clinical Psychiatry, 47, 81-82.
- Resnick, P. (1988). Malingering of posttraumatic disorders. In R. Rogers (Ed.), Clinical assessment of malingering and deception (pp. 309-327). New York: Guilford Press.
- Rogers, R. (1988). Researching dissimulation. In R. Rogers (Ed.), Clinical assessment of malingering and deception (pp. 309-327). New York: Guilford Press.
- Rogers, R. (1990). Models of feigned mental illness. Professional Psychology: Research and Practice, 21, 182-188.
- Rogers, R. (1997). Researching dissimulation. In R. Rogers (Ed.), Clinical assessment of malingering and deception (2nd Ed.) (pp. 398-426). New York: Guilford Press.
- Rogers, R., Bagby, R., & Chakraborty, D. (1993). Feigning schizophrenic disorders on the MMPI-2: Detection of coached simulators. Journal of Personality Assessment, 60, 215-226.
- Rogers, R., Sewell, K., & Salekin, R. (1994). A meta-analysis of malingering on the MMPI-2. Assessment, 1, 227-237.
- Schipper, J. (1993). Discrimination among depressive disorders using the Minnesota Multiphasic Personality Inventory - 2. (Doctoral dissertation, Texas Tech University of Kentucky, 1993). Dissertation Abstracts International, 54, 2222.
- Schretlen, D. (1988). The use of psychological tests to identify malingered symptoms of mental disorder. Clinical Psychology Review, 8, 451-476.
- Schretlen, D., Van Gorp, W., Wilkins, S., & Bobholz, J. (1992). Cross-validation of a psychological test battery to detect faked insanity. Psychological Assessment, 4, 77-83.

- Sivec, H., Hilsenroth, M., & Lynn, S. (1995). Impact of simulating borderline personality disorder on the MMPI-2: A costs-benefits model employing base rates. Journal of Personality Assessment, 64, 295-311.
- Sivec, H., Lynn, S., & Garske, J. (1994). The effect of somatoform disorder and paranoid psychotic role-related dissimulations as a response set on the MMPI-2. Assessment, 1, 69-81.
- Statistica for the Macintosh [Computer software manual], Volume 2. (1994) Tulsa, Oklahoma: Statsoft, Inc.
- Timbrook, R., Graham, J., Keiller, S., & Watts, D. (1993). Comparison of the Wiener-Harmon Subtle-Obvious Scales and the standard validity scales in detecting valid and invalid MMPI-2 profiles. Psychological Assessment, 5, 53-61.
- Wetter, M., Baer, R., Berry, D., Reynolds, S. (1994). The effect of symptom information on faking on the MMPI-2. Assessment, 1, 199-207.
- Wetter, M., Baer, R., Berry, D., Robison, L., & Sumpter, J. (1993). MMPI-2 profiles of motivated fakers given specific symptom information: A comparison to matched patients. Psychological Assessment, 5, 317-323.
- Wetter, M., Baer, R., Berry, D., Smith, G., & Larsen, L. (1992). Sensitivity of the MMPI-2 validity scales to random responding and malingering. Psychological Assessment, 4, 369-374.
- Wetter, M. & Deitsch, S. (1996). Faking specific disorders and temporal response consistency on the MMPI-2. Psychological Assessment, 8, 39-47.
- Wetzler, S., Marlowe, D.B., & Sanderson, W.C. (1994). Assessment of depression: Using the MMPI, Millon, and Millon-II. Psychological Reports, 75, 755-768.
- Ziskin, J., & Faust, D. (1988). Coping with psychiatric and psychological testimony (Vol. 2, 4th ed.). Marina del Rey, CA: Law and Psychology Press.

APPENDIX A
EXTENDED LITERATURE REVIEW

APPENDIX A

EXTENDED LITERATURE REVIEW

This review will briefly discuss the concept of malingering and the common settings and situations in which people are most likely to malingering. After ideas are presented on how real-life malingerers prepare to fake mental disorders on psychological tests, an overview of MMPI/MMPI-2 validity scales used to detect malingering will be presented. Next, MMPI malingering studies will be reviewed, followed by a more detailed literature review of similar research on the MMPI-2. MMPI-2 malingering studies that have examined the effect of specific types of information on the malingering of specific psychological disorders will be emphasized. Finally, a rationale for the present study will be provided.

The DSM-IV (1994) defines malingering as the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives such as avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs (p. 683).

DSM-IV diagnoses for factitious disorders and somatoform disorders are differentiated from acts of malingering. Factitious disorders are diagnosed when a person intentionally fabricates physical or psychological symptoms in order to satisfy a psychological need to assume the sick role (DSM-IV, p. 471). External incentives for behavior are absent in factitious disorders, whereas in malingering, external motivations are clearly evident when the circumstances are examined. Although the physical complaints typical of somatoform disorders are not fully attributable to a medical condition, the somatoform disorders differ from both malingering and factitious disorders in that the symptoms are not intentionally

produced or feigned (see Cunniën, 1988, for a more in-depth discussion).

Resnick (1988) distinguishes between pure malingering, in which a person completely feigns a disease or disorder, partial malingering, in which a person consciously exaggerates existing symptoms or falsely claims that previously legitimate symptoms are still present, and false imputation, in which a person attributes existing symptoms to a cause consciously known to have no connection to the symptoms (e.g., falsely attributing back pain to an automobile accident).

Faking mental illness on psychological evaluations given in forensic settings is particularly tempting for the potential malingerer because the rewards and benefits are so high. Persons who have been through a traumatic injury or even a minor injury have powerful incentives to either make up psychological symptoms, to exaggerate existing psychological disturbances, or to attribute existing symptoms to unrelated causes. Workers' compensation or personal injury lawsuits often involve litigants with genuine disorders such as posttraumatic stress disorder (PTSD), major depression, generalized anxiety disorder, and mild brain injury (Silvain, 1992; as cited in Lees-Haley & Dunn, 1994). Although these psychiatric disorders can legitimately result from many types of trauma (Resnick, 1988), it is not hard to imagine that some people overreport or fabricate symptoms in hopes of gaining very large financial awards.

Lees-Haley and Dunn (1994) suggest using caution in assessing self-reported psychological symptoms not only in forensic cases, but also in other settings in which the demonstration of mental illness can result in financial or other rewards. For example, Vietnam veterans with PTSD can receive financial compensation, admission to a VA hospital, or reduction of punishment for criminal behavior (Resnick,

1988). Not surprisingly, fraudulent PTSD claims are made by some Vietnam veterans. Military settings also provide powerful incentives for malingering (e.g., medical discharges and noncombat duty) (Rogers, 1990). Many people with debilitating mental disorders receive Supplemental Security Income (SSI) disability awards, commonly with diagnoses of major depression, schizophrenia, or borderline personality disorder (Reich, 1986). It is likely that some individuals have feigned mental illness in order to obtain SSI benefits. Other people may be motivated to malingering to receive non-financial help such as psychotherapy (Sivec et al., 1995). For example, if a person seeking help was judged to be in crisis or severely disturbed, he or she might receive immediate psychological help rather than being put on a waiting list.

How do malingerers prepare to fake? Persons who fabricate mental disorders may have relatively accurate and specific information about the symptoms they need to report (Wetter et al., 1993). This information can be obtained in several different ways. Some potential malingerers may have first or second-hand experience with specific disorders, either having grown up observing family members or friends who have a mental disorder or previously having had a mental disorder themselves. Most people have general knowledge of specific mental disorders through popular magazines, newspapers, television talk shows, and books (Lees-Haley & Dunn, 1994). Even television advertisements educate the public about mental illness, with some commercial psychiatric hospitals enumerating the symptoms of various psychological disorders in their advertisements. Lees-Haley (1992) reported that Los Angeles radio, television, magazine and flyer advertisements in many languages "solicit people to file personal injury and workers' compensation claims" (p. 682). These advertisements detail the symptoms that are necessary to report when filing a claim (Lees-Haley, 1992).

Besides learning about specific symptom information through the popular media, clients undergoing psychological evaluations may figure out what symptoms they are expected to report by the questions that attorneys or mental health professionals ask them (Ziskin & Faust 1988). Specifically, Lees-Haley (1985; as cited in Ziskin & Faust, 1988) reported that lawyers' questioning "defines the issues and teaches a client the range of subjects likely to arise in the case" (p. 851). Scrupulous lawyers do not deliberately advise litigants to malingering psychopathology, but

the process of exploring possible psychological disorders can create a mental set in the litigant which encourages him to falsify or exaggerate symptoms . . . questioning may even cue him as to the kinds of symptoms that might be to his advantage. (Ziskin & Faust, 1988, p.851)

A recent survey of attorneys (Wetter & Corrigan, 1994 as reported by Baer et al., 1995) revealed that a majority believed that attorneys should discuss upcoming psychological evaluations with their clients. Almost 50% of this sample of attorneys reported that attorneys should inform their clients about the presence of validity scales on psychological tests.

Besides obtaining symptom information passively or inadvertently, some sophisticated malingerers may actually prepare to fake by actively seeking out relevant symptom information from psychological textbooks and DSM-IV (Berry, 1995), or by learning about the test itself (Hare, 1985).

The MMPI/MMPI-2 is one of the most widely used personality tests used in psychological assessment (Lubin, Larsen, & Matarazzo, 1984; Butcher, 1990). One of the reasons for the widespread acceptance and use of the MMPI/MMPI-2 is the test's ability to identify major patterns of mental disorder, such as depression and schizophrenia, while also detecting troublesome test taker approaches. In particular, the MMPI and the MMPI-2 have been successful in

detecting persons who "fake bad" or malingering on the test (Berry et al., 1991; Rogers et al., 1994). Therefore, the MMPI/MMPI-2 is often used in psychological assessment batteries in settings where the possibility of malingering is a concern. For example, the MMPI/MMPI-2 is used in forensic evaluations, correctional facilities, in disability eligibility determination, and in other situations in which it might be advantageous for the test taker to feign mental illness. The MMPI's popularity as an objective personality measure, particularly in these settings, is due to its ability to detect malingering. This detection is accomplished by numerous validity scales and other indices that are designed to help identify malingerers.

The MMPI and the MMPI-2 validity scales are designed to identify inconsistent responding and to assess the accuracy of item endorsement. Because malingering is a type of inaccurate response set, much of the following discussion will focus on the scales and indices that are used for the detection of inaccurate responding on the MMPI and MMPI-2. The scales used to assess inconsistent response sets will also be examined, but in less detail. The revision of the MMPI will be briefly discussed, with an emphasis on the revision's effect on the MMPI-2's validity scales and indices.

MMPI Validity Scales

The F scale is the most widely used and most effective indicator of overreporting and malingering (Berry et al., 1991). On the original MMPI, the F scale consisted of 64 items that were answered in the scored direction by no more than 10% of the normative sample (Greene, 1991). Therefore, F scale elevations result from endorsing an unusually high number of bizarre and severe symptoms. There are multiple causes of extreme F scale elevations: (1) actual psychopathology, (2) inconsistent or random responding, and

(3) malingering (Greene, 1991). These causes may not be mutually exclusive. If an F scale elevation occurred in an MMPI record, other MMPI validity indices such as the Test Retest Index (TR) and the Carelessness Scale (CLS) could be examined to rule out random responding (a more detailed description of these scales is provided below).

Another index used in the detection of malingering is a combination of the F and K validity scales. The K scale is recognized as an indicator of defensiveness and denial of psychological distress (Berry, et al., 1991). As reported in Greene (1991), Gough (1947) suggested that subtracting the raw K scale score from the raw F scale score might aid in the detection of malingerers. Essentially, the F-K index compares the test taker's tendency to exaggerate and his or her tendency to minimize psychological problems (Berry et al., 1991). Malingering is suspected when the F scale is significantly elevated (exaggeration) and the K scale is unusually low (absence of normal defensiveness). The F-K index frequently misidentifies persons in psychological distress as overreporting. Therefore, Greene (1991) suggests using other indices and scales sensitive to malingering (e.g., F scale or Ds scale) along with the F-K index.

Another scale developed to assess malingering on the MMPI was the Gough Dissimulation Scale (Ds). This scale consists of 74 empirically-derived items that reflect stereotypes about neuroticism. The Ds scale was formed by compiling items that were answered in an opposite manner by neurotic patients and dissimulators (college students and psychologists) who were asked to fake neuroticism on the MMPI (Lamb, 1992). In 1957, the scale was revised (Ds-r) (Gough, 1957; as cited in Greene, 1991) to include only the items (40 items) that most effectively discriminated between faked and bona fide MMPI profiles (Greene, 1991). The Ds and Ds-r scales tend to be elevated by persons attempting to

feign mental illness rather than those who genuinely have a neurotic spectrum disorder (Greene, 1991).

The Wiener and Harmon Obvious and Subtle (O-S) subscales are another approach used in the detection of overreporting on the MMPI and the MMPI-2. Differentiation of obvious versus subtle item types was determined by the degree to which an item seems to tap psychopathology in a face valid manner (Wiener, 1948; as cited in Berry et al., 1991). MMPI items that seemed to explicitly indicate emotional disturbance were deemed obvious whereas items that were somewhat difficult to discern as reflecting psychopathology were categorized as subtle items (Greene, 1991). Only five scales on the MMPI/MMPI-2 contained enough subtle items to develop obvious and subtle subscales: Scales 2 (D), 3 (Hy), 4 (Pd), 6 (Pa) and 9 (Ma) (Greene, 1991). To assess for malingering using the O-S subscales, the T-score reflecting the number of endorsed subtle items is compared to the T-score reflecting the number of endorsed obvious items. When the T-score for the obvious items is much higher than the T-score for the subtle items, malingering should be considered.

There has been some research suggesting that the O-S index is of limited utility due to the invalidity of the subtle items. Graham (1993) reported that both persons instructed to exaggerate psychopathology and persons with genuine psychopathology tend to endorse many more obvious than subtle items since there are more obvious than subtle items for all scales except for scale 9. Graham (1993) concluded that "the differential endorsement of the subtle and obvious items is not particularly useful in differentiating between exaggerated profiles and valid profiles indicating severe psychopathology."

A less commonly used method of detecting malingering is Lachar and Wrobel's (1979) critical MMPI items. These 111 items are face-valid indicators of psychopathology when

answered in the deviant direction. A person attempting to overreport or malingering on the MMPI or MMPI-2 would be expected to endorse a large number of these items since they are obviously reflective of psychopathology. Greene (1991) reports that normal individuals endorse 16% to 23% of these critical items whereas psychiatric patients endorse 35% to 40% of the items. Malingerers tend to endorse an even higher percentage of the critical items than psychiatric patients.

Consistency of Item Endorsement on the MMPI

Before assessing for malingering or other types of inaccurate endorsement patterns, it is essential to assure that the test taker has answered the MMPI items in a consistent manner. For example, because elevated F scale scores may indicate random responding rather than malingering, it is imperative to rule out inconsistent responding before considering malingering. Inconsistent responding may be reflective of inadequate reading ability or comprehension, lack of educational opportunity, limited intellectual ability, excessive psychiatric or neuropsychological confusion, inability to sustain attention, substance abuse toxicity, and lack of compliance or cooperation (Greene, 1991).

Two scales were developed to detect inconsistent responding on the MMPI: the Test-Retest (TR) index and the Carelessness Scale (CLS). The TR index tallied the number of inconsistent responses to the sixteen repeated items on the MMPI (Dahlstrom et al., 1972). The 12 pairs of CLS scale items were chosen due to their opposite psychological content (Greene, 1978). A maximum total score of 12 on the CLS scale indicated that the test taker had responded to all 12 pairs of CLS items inconsistently. This scale was more subtle than the TR index in that its items were not repeated items, but were psychologically opposite and thus, more

difficult for the sophisticated client to recognize. Because both the TR index and the CLS scale consisted of relatively few items, the reliability of each scale in detecting inconsistent responding was somewhat low. Greene (1991) suggested totalling the scores on the two scales for a more reliable measure of inconsistent responding.

MMPI Revision: The MMPI-2 Validity Indices and Scales

In 1989, the MMPI was restandardized "to provide current norms, develop a nationally representative and larger normative sample, provide appropriate representation of minority groups, and update item content where needed" (Greene, 1991). In order to maintain continuity between the MMPI and MMPI-2, the validity and clinical scales of the MMPI-2 used the same criterion groups and item derivation procedures as the original MMPI (Greene, 1991). In addition to 14 new content scales and a number of supplementary scales, the MMPI-2 features several new methods of detecting inconsistent and inaccurate response sets. Furthermore, the existing validity scales were slightly modified.

On the MMPI-2, four objectionable items were deleted from the F scale leaving 60 items on the scale, and a number of F scale items were reworded (Berry, 1995). In addition, a new 40-item scale, Back F (Fb), was developed to detect infrequent responding to items appearing in the last half of the test booklet. Similar to the F scale, Fb items were endorsed in the deviant direction by no more than 10% of the normative sample. Elevations on the F and Fb scales can be indicative of malingering, severe psychopathology, or inconsistent responding. The Fb scale can be used to assess overreporting and to detect random responding (Greene, 1991). Berry (1995) recommended further validation of the scale prior to accepting it as completely analogous to the F scale for overreporting. The Fb scale is used in tandem

with the F scale to provide a measure of inconsistent responding. Specifically, since both scales are comprised of infrequently endorsed items, test takers are expected to endorse roughly the same percentage of the items on each scale. If there is a significant difference between the number of items endorsed on F and Fb ($|F-Fb|$), inconsistent responding may have occurred (Greene, 1991).

Although the F scale has been found to be the single most reliable MMPI and MMPI-2 scale in identifying malingerers (Berry et al., 1991; Rogers et al., 1994), F and Fb elevations can be produced not only by malingerers and random responders, but also by persons with severe psychological distress or severe psychopathology (Arbisi & Ben-Porath, in press). Fortunately, the MMPI-2 VRIN scale (described in more detail below) and $|F-Fb|$ can be used to detect random responders (Berry et al., 1991; Greene, 1991). However, differentiating psychopathology from faking on the basis of F elevations remains a problem in settings where there is a high base rate of psychopathology (Arbisi & Ben-Porath, in press). To minimize the confounding of psychopathology and malingering, Arbisi and Ben-Porath (in press) recently developed the F-Psychopathology scale, F(p). The F(p) scale consists of items that are infrequently endorsed by both clinical and non-clinical samples. A few studies have examined the utility of the F(p) scale in detecting malingering with very promising results (Wetter et al., 1994; Wetter & Deitsch, in press; Arbisi & Ben-Porath, in press).

Although 58 Ds scale items remain on the MMPI-2 (16 items were deleted) (Wetter et al., 1993), this scale is not scored on the MMPI-2. Therefore, there are no empirically derived scales on the MMPI-2 to detect overreporting (Greene, 1991). Even though the Ds and Ds-r scales are not officially scored on the MMPI-2, researchers are investigating the efficacy of the remaining items on these

scales in detecting malingering (e.g., Rogers et al., 1993). Lechart (1994) compared the Ds-r scale on the MMPI (40 items) to the remaining 29 items of the Ds-r scale on the MMPI-2 and found highly significant correlations (.91). Thus, the remaining Ds-r scale items on the MMPI-2 may in fact be useful as a malingering index.

The revision of the MMPI-2 led to several significant changes in the methods used to detect inconsistent responding. In an apparent attempt to shorten the test, the items for the TR index and the CLS scale were eliminated from the MMPI-2. In place of these indices, several new scales were developed to measure consistency of item endorsement on the MMPI-2. The Variable Response Inconsistency scale (VRIN) contains 67 pairs of items that have similar or opposite content (Greene, 1991). Persons who answer the MMPI-2 in a consistent manner will respond in the same direction to similar pairs and in the opposite direction to pairs of items with opposite content. If the test taker is inconsistent in his or her response to a pair of VRIN items, the pair of items is scored (Greene, 1991). According to Greene (1991), raw scores of 16 or higher on VRIN indicate that the test taker may have responded to the MMPI-2 in an inconsistent manner. He suggested using the |F-Fb| index in cases where the VRIN score is moderately elevated (raw score 8 to 15). Various combinations of VRIN and |F-Fb| used in measuring consistency of item endorsement are described by Greene (1991).

The True Response Inconsistency Scale (TRIN) contains 23 pairs of items with opposite item content. In contrast to the VRIN scale, the TRIN scale is not a measure of inconsistent responding. Rather, it is designed to detect true or false response sets. It also differs from the VRIN scale in that the scored response for both items in the pair is either "true" or "false" (Greene, 1991). For instance, if a test taker answers both items 99 and 314 "true," the

pair is scored as an contradictory or inconsistent response (Greene, 1991). For a more thorough discussion of TRIN scoring procedures, see Greene (1991). High TRIN scores are reflective of clients who tend to answer "true" and low scores are produced by clients who tend to answer "false" regardless of the item content (Greene, 1991).

Research on the MMPI and Malingering

Many studies have assessed the effectiveness of the previously mentioned scales and indices to identify response sets on the MMPI and MMPI-2. Because this study will examine the MMPI-2 and malingering, this literature review will focus mostly on relevant MMPI-2 studies on malingering. However, a general and brief review of earlier MMPI malingering studies will also be provided.

The following summary of MMPI malingering studies is based on the meta-analytic findings of Berry et al., (1991). The meta-analysis included 28 MMPI malingering studies (1947-1989) with a total of 44 group comparisons. When comparing honest versus faking groups on F, F-K, Ds, Ds-r, and O-S, the mean effect size was 2.07 overall. Considering that Cohen (1977; as cited in Berry et al., 1991) recommended interpreting a difference score of .80 or higher as a large effect, 2.07 is a particularly large effect. When contrasting faking and honest groups, the largest effect sizes were found for the F scale (2.34), the F-K index (1.89), and the Ds scale (2.17). Because the meta-analysis indicated that Ds-r and O-S were significantly less effective in detection of malingering, Berry et al. (1991) suggested using caution in using those two measures to assess for malingering.

Berry et al. (1991) reported that the effect sizes based on comparisons of faking and honest groups were influenced by methodological variables including the content of instructions and the type of design and sample used in

the study (see Table A.1). Instructions ranged from asking participants to fake mental illness in a global sense (e.g., Grow, McVaugh, & Eno, 1980) to instructions requesting participants to fake a specific disorder such as closed head injury (e.g., Heaton, Smith, Lehman, & Vogt, 1978). Although both kinds of instructions produced detectable malingering profiles, groups using global fake bad instructions were even more easily detected. In general, participants asked to fake mental disorders in a general sense on the MMPI are detected by their propensity to endorse large numbers of items that are conspicuously pathological or are infrequently endorsed by patients who actually have mental disorders (Wetter et al., 1993). Five of the 28 studies from the meta-analysis compared hypnotized participants with non-hypnotized participants in their ability to successfully mangle on the MMPI. Although results showed that hypnotized participants were less easily detected, Berry et al. (1991) noted that "clinicians may be reassured upon reflecting that few of their patients present themselves for evaluation in a hypnotic trance" (p. 596).

Another methodological difference that appears to affect the degree of success in malingering performance is the type of research design used. Berry et al. (1991) found that malingering studies on the MMPI employing repeated measures designs tended to produce larger effect sizes than did between groups designs (2.38 vs. 1.92). However, taking into account that the "clinician's task is to classify different patients into honest or faking groups rather than different protocols from the same patient into honest or faking categories," it would appear that between groups designs have greater external validity (Berry et al., 1991, p. 596).

Berry et al. (1991) also found that different types of samples are able to mangle more successfully (i.e., are able to elude detection by the validity scales). Among the

44 comparisons examined in the meta-analysis, comparisons included college students (21), military personnel (9), patients (6), prisoners (5), and disability applicants (3). The largest effect size resulted when normals (such as college students and military personnel) answering honestly were compared to normals faking bad (2.66). Lower, but still impressive effect sizes were found for the following two comparisons: (1) normals faking bad compared to patients answering honestly (1.86), and (2) patients answering honestly versus patients faking bad (1.48). The lowest effect size (0.83) compared participants inferred to be malingering to various other groups.

Overall, these results demonstrate that as the ecological validity of the comparison increased, effect size decreased (Berry, 1995). For example, more clinically relevant comparisons such as normals faking bad vs. patients answering honestly, or patients answering honestly vs. patients exaggerating, tended to produce smaller, but significant effect sizes when compared to comparisons involving college students. Although malingering studies using student samples have excellent internal validity, results may not be generalizable to other more relevant populations such as disability applicants, personal injury claimants, or criminals, all of whom presumably have a higher base rate of malingering.

Rogers (1988) and Schretlen (1988) have identified several methodological problems in malingering research. Noting that the instructions used were often "brief and nonspecific," Rogers (1988) suggested that instructions to research participants who are attempting to malingering should emphasize the presentation of a believable disorder. Vague instructions (e.g., "appear as mentally ill as you can") may influence participants to make no effort to appear credible (Rogers, 1988). Rogers (1988) suggested dealing with this problem by either drafting instructions that emphasize

appearing believable or by providing a detailed description of a hypothetical situation (i.e., a scenario). In any case, participants should be made aware of the importance of not over-dramatizing the mental illness they have been instructed to fake (Rogers, 1988). Giving participants more preparation time may also increase participants' level of involvement and motivation (Rogers, 1988).

Lack of debriefing is another problem in malingering research. Debriefing can be used as a manipulation check to gauge participants' comprehension of and compliance with instructions and can also serve as a way of exploring the various ways that participants attempt to malingering (Rogers, 1988).

Both Schretlen (1988) and Rogers (1988) pointed out the weaknesses of pure analog designs. They suggested the use of psychiatric control groups or other groups (e.g., prison inmates) to increase the clinical relevance of malingering research.

Most MMPI research has not offered rewards to participants for effective malingering despite the fact that actual malingerers are almost solely motivated by real or perceived external rewards (Rogers, 1988). Both authors stressed the need for offering external incentives to participants for successful faking in order to increase external validity (Rogers, 1988; Schretlen, 1988). Many studies focusing on the MMPI-2 and malingering have taken steps to address the key methodological problems in malingering research identified by Schretlen (1988) and Rogers (1988).

Research on the MMPI-2 and Malingering

The following review of MMPI-2 malingering studies will focus first on "fake bad" studies. "Fake good" comparisons will be excluded from examination. Then, a more detailed review will be provided for MMPI-2 malingering studies that

instruct participants to fake specific mental disorders and those that "coach" participants on symptom information and MMPI-2 validity scales.

Fake-bad studies

The majority of studies on malingering on the MMPI-2 are similar to those with the MMPI in employing analog designs using college students who are instructed to "fake bad" or malingering general mental illness. Three MMPI-2 malingering studies have used college students in all groups and have lacked patient comparison groups (Austin, 1992; Cassisi & Workman, 1992; Wetter, Baer, Berry, Smith, & Larsen, 1992).

Austin (1992) compared students instructed to fake bad with students answering the MMPI-2 under standard instructions. Taking into account Rogers's (1988) recommendations about instructions, Austin (1992) provided fake bad participants with a brief scenario that stressed believability of presentation. The specific instructions were to: "respond as though you are mentally ill and trying to increase your chances of getting disability benefits or of beating a criminal charge. Be as clever as possible and try to answer in such a way that it will not be evident you are faking" (Austin, 1992, p. 670). Results indicated that several MMPI-2 validity indices (F, F-K, K) were highly accurate in identifying fake bad participants, with the obvious-subtle (O-S) scale being somewhat less effective. Specifically, Austin (1992) found that raw scores of F>29, F-K>11, and K<31 each, on their own, accurately identified at least 95% of fake bad participants, whereas O-S>190 correctly identified 84% of the fake bad participants (Austin, 1992).

Using a truncated version of the MMPI-2 consisting of only the 102 validity scale items, Cassisi and Workman (1992) compared college students asked to fake bad with

those asked to answer honestly. Similar to the scenario used by Austin (1992), Cassisi and Workman's (1992) instructions stressed believability of presentation and provided a rationale for faking a mental disorder. Using the standard rule for detecting malingering (a T-score for scale F of 70 or greater; Butcher et al., 1989), Cassisi and Workman (1992) accurately identified 95% of the fake bad group. Ten percent of the group answering the MMPI-2 under standard instructions were falsely identified as faking bad using the F scale greater than or equal to 70 cutting score (Cassisi & Workman, 1992).

Wetter et al. (1992) compared college students in four conditions: fake severe psychological disturbance, fake moderate psychological disturbance, answer randomly, and answer under standard instructions. The fake moderate group received instructions that read in part, "Please complete the test as if you were trying to create a bad impression. Imagine that you are attempting to appear sufficiently disturbed so as to obtain money for your psychological problems, such as in a lawsuit against a person or company you claim hurt you psychologically or for mental disability from Social Security funds. At the same time, however, try not to appear so severely disturbed that psychiatric hospitalization would be required." In contrast, instructions for the fake bad group read, "Please complete the test as if you were trying to create a very bad impression. Imagine that you are attempting to appear sufficiently disturbed so as to appear as if you were not responsible for your own actions, such as avoiding prosecution for a crime or attempting to get into a psychiatric ward for some reason. In other words, we want you to fake serious mental illness." Thus, an emphasis was placed on faking in a believable manner in the fake moderate group, while the fake bad group was encouraged to appear as disturbed as possible. Results showed that random and

malingered responses produced significant elevations on F, Fb, and Ds-2 (items from the MMPI Ds scale that remain on the MMPI-2), whereas only random responders elevated VRIN. Results indicated that the F-K index and Ds-2 scale elevations increased significantly as the degree of faked psychopathology increased from moderate to severe (Wetter et al., 1992). The authors did not establish optimal cutting scores or report percentages of correct identification.

Three other fake bad studies have compared college student participants to genuinely disturbed psychiatric patients in an effort to increase clinical relevance (Bagby et al., 1994; Graham, Watts, & Timbrook, 1991; Timbrook, Graham, Keiller, & Watts, 1993). In one study, Graham et al. (1991) used a repeated measures design to examine the effectiveness of the MMPI-2 in detecting fake good and fake bad responding by college students. Specifically, college students were asked to complete the MMPI-2 two times, once with standard instructions, and again with instructions to either fake good or fake bad. Fake bad participants' profiles were compared to the psychiatric inpatient profiles and to the profiles of college students who had completed the MMPI-2 with standard instructions. Malingering instructions did not mention the importance of responding in a believable manner: "Answer in the way that you think one would answer if he or she wanted to give the impression of being a person who has very serious psychological or emotional problems." Using a cutting score of F>18, 94% of the students in the malingering condition were correctly identified, but a high rate of misclassification occurred among inpatient profiles (33% of male and 20% of female inpatients were classified as malingering using this cutting score). The F-K index also accurately discriminated between the fake bad and standard instruction profiles. For students, the most accurate classification rate resulted using an F-K raw cutting score of 12 for women (80% of

patients, 95% of fakers, and 100% of honest responders correctly identified) and 17 for men (93% of patients, 93% of fakers, and 100% of honest responders correctly identified). To accurately classify over 94% of inpatient profiles, the authors used an F-K raw score cutting score of 27 for men and 25 for women. The Fb scale was also very successful in identifying fake-bad profiles. When fake-bad profiles were compared with standard instruction profiles, Graham et al., (1991) reported different optimal raw cutting scores for men (Fb scale greater than or equal to 19 with 93% of fakers correctly identified) and women (Fb greater than or equal to 22 with 95% of fakers correctly identified). In order to correctly classify 90% of the inpatients, slightly higher Fb scale cutting scores were used (23 for men and 24 for women). The authors cautioned that using cutting scores that distinguish fake-bad profiles from honest profiles in normal populations are "inappropriate for use in a psychiatric population" (Graham et al., 1991, p. 274). In other words, optimal cutting scores in a normal population will lead to high rates of misclassification of inpatient profiles as fake bad. In addition, the authors cautioned that the use of a single validity scale score cutting score, regardless of gender, is likely to lead to higher rates of misclassification (Graham et al., 1991).

In another repeated measures design, Timbrook et al. (1993) focused on the utility of the Wiener-Harmon Subtle-Obvious scales (Wiener, 1948; as cited in Timbrook et al., 1993) in detecting fake good and fake bad responding. The study compared psychiatric patients to college students who were instructed to either fake good, fake bad, or respond honestly. The instructions given to the fake bad group were very similar to those used by Graham et al. (1991):

In responding to the items in the inventory, please try to answer in the way that you think one would answer if

he or she were wanting to give the impression of being a person who has very serious psychological or emotional problems" (p. 266-267).

Although Timbrook et al. (1993) found that the O-S index correctly identified a large percentage of fake bad profiles (O-S>190 identified 90% of male patients, 100% of male fakers, 80% of female patients, and 95% of female fakers), a hierarchical regression analyses showed the O-S index did not account for as much variance as the F scale. Timbrook et al. (1993) concluded that the F scale appeared to be more useful than the O-S scales in the detection of fake bad profiles.

To determine the magnitude of group differences, Timbrook et al. (1993) calculated effect sizes. Comparing the faking bad group and the standard instructions group, the F scale produced the largest effect sizes ($d = 1.84$ and 1.88 for men and women, respectively). The sum of the obvious scale items yielded an effect size of 1.85 for men and 1.82 for women. The O-S index produced slightly smaller effect sizes (ranging from 1.82 to 1.79). In comparisons between patients answering honestly and college students who were faking bad, F scale effect sizes ranged from 1.80 to 1.82 , sum of obvious scale items ranged from 1.65 to 1.63 , and the O-S index ranged from 1.59 to 1.52 . An O-S index score cutting score of 190 (suggested by Greene, 1991) correctly classified 90% of the male patients, 90% of male fakers, and 100% of honest responders of both genders. Although only 80% of female patients were accurately categorized with O-S index greater than or equal to 190 , 95% of females fakers were correctly classified. Interestingly, the authors discovered that the sum of the obvious scores alone were just as effective as O-S in making correct classifications, calling into question the usefulness of the subtle items in the O-S index (Timbrook, et al., 1993).

In a between-groups design, Bagby, Rogers, Buis, and Kalemka (1994) compared college students instructed to fake good, fake bad or answer honestly to psychiatric patients given standard instructions. Following the advice of Rogers (1988) and Schretlen (1988), Bagby et al. (1994) instructed fake bad participants to respond to the MMPI-2 in a believable way and offered monetary incentives to successful fakers in an attempt to increase motivation in the study. All validity scales examined correctly classified 80 to 82% of inpatients. However, large differences were seen in the classification rates of fake bad profiles: 92% ($\underline{F}>15$), 95% ($\underline{F-K}>7$), 78% ($\underline{Fb}>93$), 62% ($\underline{Ds-r2}>18$), 67% ($\underline{O-S}>179$), and 71% (Lachar and Wrobel critical items >58). Cutting scores for $\underline{F}$ and $\underline{F-K}$ were higher than those reported by previous research. For example, Graham et al. (1991) suggested a cutting $\underline{T}$ score of $\underline{F}$ greater than or equal to 107 compared to Bagby's et al. (1994) $\underline{F}>89$. The extreme differences in cutting scores for these scales may have resulted from differences in the clinical groups employed as comparison groups. Specifically, Bagby et al. (1994) mentioned that each study has had a different percentage of psychotic patients and these patients are likely to endorse the unusual symptomatology that produces $\underline{F}$ scale elevations. In addition, Bagby et al. (1994) suggested that cutting scores which are based on small numbers of participants are "likely to capitalize on chance variations" (p. 37).

Several studies have compared college students faking psychological disturbance with forensic populations (Bagby, Buis, & Nicholson, 1995; Bagby, Rogers & Buis, 1994). Forensic populations were used to increase the external validity of the findings. Bagby et al. (1994) produced the first study using a forensic sample as a comparison group. Specifically, the group consisted of pretrial, forensic inpatients. In this between groups design, college students in the fake bad condition were instructed to fake in a

believable manner and were given different descriptions of situations in which malingering might occur. To increase their motivation, participants who successfully malingered were eligible to win a \$100 lottery. Results showed that F, F-K, Ds-2, and O-S were moderately effective in detecting fake bad profiles (Bagby et al., 1994). Specifically, correct classification of inpatients ranged from 80 to 82% using the cutting scores for the scales and indices listed below.¹¹ In descending order, correct classification of faked profiles for various validity scales were 88% for F>104 and F-K>11, 83% for Fb, 81% for Ds-r2, 73% for O-S>169, and 71% for Lachar and Wroebel critical items >57 (Bagby et al., 1994).

Using a similar method, Bagby et al. (1995) compared forensic inpatients and general psychiatric patients to college students instructed to fake bad. The two goals of the study were to corroborate the results of Timbrook et al. (1993), who found that the O-S subscales were not particularly helpful in comparison to the F scale in detecting malingering and to test the utility of the Ds-r2 scale (MMPI-2 items left from the MMPI Ds-r scale) in identifying fakers. Bagby et al. (1995) found that the F scale was the best predictor for detecting students who were faking bad. For comparisons between fake bad participants and inpatients, the effect size for the F scale was 2.47 for men and 2.61 for women. Effect sizes for the Ds-r2 scale ranged from 1.34 (men) to 1.47 (women). The smallest effect sizes were found for the O-S index (1.15 for men, 1.49 for women) and the Obvious items sum (0.93 for men, 1.10 for women). Similar effect sizes were found for the validity scales in comparisons between forensic inpatients and the

¹¹In the Bagby et al. (1994) study, the authors appeared to assume that the forensic inpatient group consisted of non-malingering individuals.

fake bad group. Neither the O-S subscales or the Ds-r2 scale were found to have sufficient incremental validity beyond the F scale to be worthwhile in the prediction of malingering. Therefore, the authors cautioned against relying too much on the use of the O-S subscales in the assessment of fake bad responding (Bagby et al., 1995).

Only one study to date has used prison inmates in an experimental malingering design. In a between-groups design, Iverson, Frazen, and Hammond (1995) examined inmates' ability to malingering on the MMPI-2 by instructing them to fake serious mental illness without being obvious that they were faking. The scenarios used were geared toward a prison population:

You have committed a felony and are awaiting arraignment... You decide to fake mental illness in order to influence the process. You are not sure if it will work, but you figure that if you are convicted you might get probation or a lighter sentence if the judge thinks you have serious mental problems.

No incentives were used to motivate inmates to produce an undetectable malingered profile. Experimental malingerers were compared to psychiatric inpatients and inmates who were instructed to respond honestly. Iverson et al. (1995) reported that the F scale was more accurate than the Fb scale in identifying fakers. Specifically, F>17 correctly identified 89% of faking inmates and 98% of inpatients. In comparison, Fb>17 correctly classified 61% of faking inmates and 96% of inpatients. Various cutting scores were reported for the F-K index (ranging from 9 to 12) with correct classification of 98% of inpatients and greater than 82% of faking inmates. Overall, the F scale and the F-K index were found to provide slightly better overall categorization rates than did Fb (Iverson et al., 1995).

Lees-Haley (1991a, 1991b, 1992) has conducted an impressive line of research on personal injury claimants. His MMPI-2 malingering studies are different from previous

research (including the present study) because he used suspected malingerers as participants. Rather than instructing analog participants to fake bad, Fox, Gerson, and Lees-Haley (1995) examined the MMPI-2 profiles of claimants who had previously taken the MMPI-2 as a part of a comprehensive evaluation related to a personal injury claim.

Some of these studies compared non-malingering personal injury claimants with claimants who were obviously malingering (Lees-Haley, 1991a; Lees-Haley, 1991b; Lees-Haley, 1992). Criteria used to determine presence of malingering was "convincing, concrete evidence of misrepresentation of functioning and gross discrepancies between self-report and objective functioning" (Lees-Haley, 1991a, p. 816). In general, results have shown that F and F-K have utility for differentiating malingering vs. nonmalingering personal injury claimants. However, Lees-Haley (1991b, 1992) suggested using lower cutting scores for forensic populations than those suggested by other MMPI research. Lees-Haley (1991b) cited an earlier study (Lees-Haley, 1989) that explains that personal injury malingerers tend to obtain lower F and F-K scores than criminals faking insanity because some of the answers needed to produce high F and F-K scores require admissions that personal injury malingerers are reluctant to make. In other words, personal injury malingerers tend to simultaneously fake good and fake bad. They present themselves as sincere and "relatively normal with the exception of the problems associated with their injury" (p. 7).

Although Lees-Haley's quasi-experimental studies are unlike most other MMPI and MMPI-2 research on malingering, they are an important contribution to this area of research. His use of actual malingerers as participants increases external validity because they are a naturally occurring group who have been motivated to fake on the MMPI and MMPI-2 in order to obtain substantial real-life rewards.

Studies providing symptom information

Upon the recommendation of researchers (e.g., Berry et al., 1991; Schretlen, 1988) to provide specific symptom information to participants who are asked to fake mental disorders, a new line of research on malingering on the MMPI-2 has emerged in recent years. Researchers have "coached" participants by providing them with symptom information about the particular disorders they have been asked to fake (Bagby et al., 1997; Lamb et al., 1994; Rogers et al., 1993; Sivec et al., 1994, 1995; Wetter et al., 1993, 1994; Wetter & Deitsch, 1996). These studies have taken into account the possibility that some malingerers may actively seek out information on disorders or may be given the information either overtly or unintentionally (Lees-Haley, 1992; Ziskin & Faust, 1988). Clearly, information on specific disorders is widely available to the lay person through the media.

Wetter et al. (1993) conducted one of the first studies to examine the effect of symptom information on participants' ability to fake specific disorders on the MMPI-2. Normal adult participants were provided with information on posttraumatic stress disorder (PTSD) and paranoid schizophrenia, and then were instructed to simulate these disorders on the MMPI-2. Symptom information was compiled from magazine articles and pamphlets that were easily accessible to the public. Participants in the faking conditions were compared to patients with actual PTSD or paranoid schizophrenia to determine if symptom information helped participants to more believably fake the disorders on the MMPI-2. To increase external validity, Wetter et al. (1993) followed the suggestions of Rogers (1988) and Schretlen, Van Gorp, Wilkins, and Bobholz (1992) by offering financial incentives for successful faking. To increase participant involvement and compliance in the study, participants were given understandable information about the

disorders, were required to fill out a worksheet on the disorder, and were quizzed on the symptom information. A feedback questionnaire was also administered to make sure that participants understood and complied with the instructions.

Results showed that fakers obtained higher scores on F, Fb, F-K, Ds, and all of the clinical scales in comparison to patients (Wetter et al., 1993). Specifically, a T-score greater than 96 for the F scale correctly classified 85% of normals faking PTSD and 70% of PTSD patients, while F>104 correctly identified 100% of paranoid schizophrenia fakers and 90% of patients with the disorder. F-K>15 and Ds-2>35 each separately correctly classified 75% of PTSD fakers and 85% of PTSD patients. Fb (>98) had the lowest hit rate among PTSD groups, correctly identifying 80% of fakers, but only 65% of patients. For the paranoid schizophrenic groups, Ds-2 proved to be the weakest validity scale, correctly identifying 86% of fakers and 95% of patients (Wetter et al., 1993).

To determine the magnitude of the effects, effect sizes for the validity scales were calculated. Comparisons between PTSD fakers and patients produced effect sizes for F, F-K, and Ds-2 that were greater than 1.00, whereas comparisons between paranoid schizophrenic fakers and patients yielded effect sizes that were greater than 2.00 (Wetter et al., 1993). The authors concluded that having access to specific information about a particular disorder may not help malingerers to generate profiles which closely resemble those produced by patients with the disorder or to elude detection by the MMPI-2 validity scales and indices (Wetter et al., 1993).

In the next study, Wetter et al. (1994) examined the effects of giving specific information about borderline personality disorder (BPD) on the malinger success of normal adults on the MMPI-2. Four groups of adults were

investigated: normal adults given specific BPD information and asked to simulate it, normal adults asked to simulate a non-specific mental disorder without specific symptom information (fake-bad group), patients with a BPD diagnosis, and normal adults who were instructed to answer the MMPI-2 honestly. This was the first study to compare demographically matched groups of patients and informed fakers with an uninformed faking group. Similar to Wetter et al. (1993), this study followed the methodological recommendations of Rogers (1988) and Schretlen (1988) by offering monetary incentives for successful faking and by using participants from the general population. In addition, participants took quizzes to test their comprehension of the symptom information and filled out debriefing questionnaires to evaluate their understanding of the disorder and the intelligibility of the material they studied.

Instructions given to the uninformed fake-bad group asked participants to "create a bad impression..by pretending to be psychologically or emotionally disturbed," emphasizing that they should appear disturbed enough to receive benefits, but not so "severely disturbed that psychiatric hospitalization would be required" (Wetter et al., 1994, p. 202). Participants asked to fake BPD on the MMPI-2 were given BPD symptom information from an article in a popular magazine. Specific instructions for this group were not published in the journal article by Wetter et al. (1994).

Results were consistent with Wetter et al. (1993) in that specific symptom information did not appear to allow participants to successfully fabricate a mental disorder on the MMPI-2. In fact, the informed fakers scored similarly to the uninformed fakers, elevating the validity scales (F, Fb, F-K, F(p)) significantly higher than patients (Wetter et al., 1994).

The authors reported effect sizes for comparisons between experimental malingerers with symptom information and patients on the F scale (1.02), Fb scale (.92), F-K index (.98), Ds-2 scale (.90) and F(p) scale (1.30). When comparing fakers with symptom information to fakers without symptom information, effect sizes ranged from .02 (F(p)) to .39 (Fb). Using a cutting T score of 98, the F scale correctly classified 70% of the BPD fakers and 81% of the BPD patients. The Fb scale (>104) correctly identified 87% of the fakers and 69% of the patients, while the F-K index (>13) correctly classified 61% of BPD fakers and 86% of patients. Using a raw Ds cutting score of 33, 61% of the fakers and 83% of the patients were accurately identified. With an overall hit rate of 83%, the newly developed F(p) scale identified 74% of fakers and 89% of patients, making it the most successful indicator of malingering in this study. The authors concluded that "symptom information appears to almost hinder rather than help fakers when they complete the MMPI-2" (Wetter et al., 1994, p. 204).

Similar to Wetter et al. (1994), Sivec et al. (1994) examined the impact of specific versus nonspecific instructional sets on malingering performance. In a repeated measures design, undergraduates first took the MMPI-2 under standard instructions and then were randomly assigned to one of three instructional groups: somatoform disorder, paranoid psychotic, and a general "fake bad" group. After taking the MMPI-2 under standard instructions, participants in the instructional conditions received information about the role they were to enact and were given 2 days to prepare for the second testing session. Monetary incentives were provided to encourage motivation.

Compared to previous research, Sivec et al. (1994) used a more specific, realistic, and lengthier scenario. The scenario involved having the participant imagine faking symptoms in a psychological evaluation to receive financial

compensation from a pharmacy that gave the participant the wrong prescription. Instructional groups were given information either about somatoform disorder based on a 1-3/3-1 codetype description, paranoid psychotic description based on a 6-8/8-6 codetype description, or a brief global fake bad description (feign "serious psychological and emotional problems"). All three instructional groups were told to portray their role in a believable manner on the MMPI-2.

A feedback questionnaire was administered to assess the time spent in preparation for the role and the degree of confidence participants felt about being able to portray the role accurately. Because few studies have attempted to understand how malingerers prepare to fake, Sivec et al. (1994) included an open-ended question ("How did you prepare for the role?") to determine simulators' preparation strategies. Results indicated that participants used a variety of strategies ranging from simply reading the scenario, to consulting a textbook, to drawing from personal experience.

Results demonstrated that the validity indices, particularly the F scale, were very successful at detecting participants in the fake bad and paranoid psychotic groups (correctly classified at rates of 93%-100%), but were less accurate in detecting the somatoform group (40%-74%). Although the somatoform group differed dramatically from the standard instruction group, none of the available validity indices or scales could identify these fakers with a high degree of accuracy. Hierarchical regression analyses revealed that the "O-S items and the T score of total obvious items contributed significant increments to the predictions based upon the F scale alone for the paranoid psychotic and the fake-bad analyses." For the somatoform group, the F scale was the best predictor of malingering, with the F-K index, the Ds-2, and the T-score of total

obvious items adding very little to accurate prediction. The authors concluded that the MMPI-2 is "highly effective" in identifying college students instructed to "fake-bad" or to fake a paranoid psychosis, but that the MMPI-2 is "less capable of discriminating those individuals who present with feigned somatoform disorder symptoms in the neurotic spectrum" (p. 79).

Since most research on malingering and the MMPI/MMPI-2 has focused on Axis I disorders, Sivec et al. (1995) decided to examine the effect of symptom information on the performance of undergraduate participants instructed to simulate borderline personality disorder (BPD). The BPD faking group was compared to a clinical comparison group of persons diagnosed with BPD, a group of participants instructed to fake major depression, and a group of honest responders. Since some authors contend that BPD is a "subaffective" disorder, the investigators included the depression faking group so that the differences between faked BPD and faked depression presentations on the MMPI-2 could be examined.

Participants in the experimental conditions (BPD and depression) were given detailed, specific scenarios that emphasized believability of performance. Descriptions of BPD and depression based on the DSM-III-R (1987) diagnostic criteria were written in the second person (e.g., for the BPD group, "You are highly impulsive and prone to engage in various forms of self-damaging behavior..."). To increase the likelihood that participants would cognitively process the role given to them, experimental groups were supplied with a "role preparation sheet" which was later examined to make sure participants accurately interpreted the instructions. Monetary incentives were provided to increase motivation and involvement in the task.

Consistent with previous research (Rogers et al. 1993; Sivec et al., 1994; Wetter et al., 1993, 1994), results

showed that participants given specific symptom information are easily detectable by the MMPI-2 validity scales. Due to overendorsement of symptoms, both BPD and depression simulators obtained significantly higher elevations on the F scale and the clinical scales than the BPD patient group and the control group (Sivec et al., 1995). The authors reported that an F scale score greater than or equal to 19 correctly identified 95% of the BPD fakers and 90% of the genuine BPD patients. A discriminant function analysis revealed that the rate of overall classification using the F scale, F-K index, Ds-2, and O-S, was 93.3%, the same outcome was accomplished by using the F scale (>19) by itself (Sivec et al., 1995).

Bagby et al. (1997) examined the ability of student fakers to malingering depression and schizophrenia on the MMPI-2 when given DSM-IV criteria for each disorder. The depression faking group (n = 20) was compared to a group of depressed patients (n = 40) and the schizophrenia faking group (n = 20) was compared to a group of schizophrenic patients (n = 40).

Results indicated that the validity indices were better at detecting faked schizophrenia than they were in detecting faked depression. Bagby et al. (1997) indicated that familiarity with depressive experiences may have caused this difference. Due to overendorsement of symptoms, both depression and schizophrenia fakers obtained significantly higher elevations on all the validity indices (F, Fb, F-K, F(p), O-S, Ds-2, and the Obvious items from the O-S index). In addition, both the depression and schizophrenia fakers obtained significantly higher elevations on a five of the clinical scales (for the depression group, scale 2, 4, 6, 8, and 0; for the schizophrenia group, scale 1, 4, 6, 8, 9, and 0) in comparison the respective clinical comparison groups.

Using a hierarchical regression analysis, F and Fb were shown to be the best validity indices at differentiating MMPI-2 profiles from depressed patients and student fakers of depression. In addition, F, Fb, and F(p) were found to be the most effective validity indices in differentiating patients with schizophrenia from students asked to fake schizophrenia on the MMPI-2.

The authors reported effect sizes for comparisons between experimental malingerers and patients with depression or schizophrenia. The largest effect sizes in differentiating depression fakers from depressed patients were for Fb (3.30) and F (3.13) whereas the largest effect sizes distinguishing schizophrenia fakers from inpatients with schizophrenia were for F (3.99), F(p) (3.77), and F-K (3.33).

Studies providing validity scale information

Although it has been a concern that malingerers may prepare for taking the MMPI/MMPI-2 by studying specific symptom clusters, research examining the effect of specific symptom information on malingered efforts have shown that this information does not help fakers elude detection on the MMPI-2. However, several authors have questioned the assumption that persons fabricating mental disorders are naive to the purposes of psychological testing (Lamb et al., 1994; Rogers et al., 1993). Some authors have proposed that some sophisticated malingerers may prepare for faking on a personality test by studying the test itself and how the test assesses malingered (Lamb et al., 1994; Rogers et al., 1993). Hare (1985), in a study of male prison inmates, describes one such instance:

One inmate had his own MMPI manual, a set of scoring keys, and several books on the clinical interpretation of the MMPI. As a result, he was able to produce a given type of profile on demand, not only for himself, but also for other inmates. For the latter, he

operated his own consulting service, advising inmates on how they should respond to certain types of items in order to obtain the "appropriate" MMPI profile. Not all the inmates were as sophisticated or as enterprising as this one was, but it is probable that many were more test-wise and in better control of the testing sessions than some clinicians and investigators might imagine (p. 9).

According to a recent study of attorneys' attitudes (Wetter & Corrigan, 1994 as cited in Baer et al., 1995), many attorneys believe that it is "their professional responsibility to inform and prepare clients" who are going to be psychologically evaluated. Almost 50% of the attorneys participating in the survey reported that they should inform their client about the presence of validity scales on psychological tests. Although it is unlikely that many lawyers "coach" their clients on methods of outsmarting the MMPI-2 validity scales, an attorney's practice of informing the client about the existence of validity scales may encourage the client to seek out more information about the MMPI-2 and the validity scales. Clearly, it is important to understand how knowledge of the MMPI-2 validity scales and detection strategies can affect malingering effectiveness.

To date, only two studies have considered the possibility that malingerers might actually prepare for taking the MMPI-2 by studying how the MMPI-2 detects faking (Lamb et al., 1994; Rogers et al., 1993). Participants in both studies were coached about the MMPI-2 validity scales to determine the effect of this information on malingering performance (Lamb et al., 1994; Rogers et al., 1993).

In a between-groups design, Rogers et al. (1993) randomly assigned 72 normal adults to one of four malingering groups: coached on schizophrenia symptoms, coached on strategies for detection of fakers, coached on both symptoms and strategies, or uncoached. These four groups of normal participants were compared to 13 honest

responders and to 37 inpatients with schizophrenia. Participants in the experimental conditions were instructed to fake a schizophrenic disorder "of sufficient severity as to warrant hospitalization" (Rogers et al., 1993). They were given the appropriate information depending on group membership and were given the MMPI-2 to take home and complete during the span of a week. In concordance with Rogers's (1988) suggestions on improving malingering research, participants were offered monetary incentives for successfully eluding detection.

The validity scale information contained strategies for the detection of faking associated with the F scale, the Ds-r scale, O-S items, and the Lachar-Wrobel critical items. The information not only involved tips for avoiding detection, but also briefly mentioned the names of the scales and how some of the scales were developed. Schizophrenia symptom information was based on DSM-III-R criteria (1987) and with technical terms such as hallucinations and catatonic behavior clearly explained. Participants were not quizzed on their retention or their level of understanding of the validity or symptom information.

Results were consistent with previous research (Wetter et al., 1993, 1994; Sivec et al., 1994, 1995) in showing that specific symptom information did not appear to help fakers escape detection. However, about 33% of the participants coached about MMPI-2 detection strategies were able to avoid detection (Rogers et al., 1993). Participants coached on both validity indices and schizophrenia symptom information were less successful than those coached on detection strategies alone. Rogers et al. (1993) hypothesized that participants given both kinds of information were distracted by the symptoms and made less effective use of the strategies for avoiding being detected by the validity scales.

The O-S index and the Ds-r2 index were the most successful validity indicators, correctly classifying 64% of those coached on strategies alone (O-S>106, Ds-r2>15). The O-S and Ds-r2 indices were also successful in classifying large percentages of those coached on symptoms only (88%, 80%, respectively) and those coached on both strategies and symptoms (93% and 73%, respectively). The F scale and the F-K index were less successful in identifying participants coached on strategies, either with symptom information or without. Of the group coached on strategies alone, only 21.4% were correctly identified by a F>28 (raw score) and only 28.6% were correctly identified by raw scores of F-K>14 and Fb>17. The F scale and the Fb scale correctly identified 73.3% of the group coached on symptoms only and 46.7% (F scale) and 40% (Fb scale) of the group coached with both types of information (Rogers et al., 1993).

In a similarly designed study, Lamb et al. (1994) examined the effects of validity and symptom information on the malingering of closed head injury (CHI) on the MMPI-2. The investigators assigned 179 undergraduates to four groups analogous to the groups used by Rogers et al. (1993). A control group consisting of 91 undergraduates completed the MMPI-2 honestly, and no clinical comparison group was used. The methodology used was similar to that used by Wetter et al. (1994) in that incentives, worksheets and quizzes on symptoms and strategies, and feedback questionnaires were used to assess levels of motivation and comprehension. The scenarios were detailed and explicit, stressing a believable simulation of CHI on the MMPI-2. Validity information included detailed and technical information about the scales as well as some uncomplicated tips for eluding detection. The quizzes on validity information primarily tested participants' retention and comprehension of information related to the functions of the scales, the names of the scales, and how they were developed rather than on MMPI-2

detection strategies. Participants who answered less than 70% of the quiz items correctly were excluded from the analysis. Feedback questionnaires were more thorough than those used by Wetter and colleagues (1993, 1994) with questions examining clarity of the scenario, perceived success in assuming the scenario role, ability to maintain the role throughout the MMPI-2, and extent of random responding (Lamb et al., 1994).

Consistent with Rogers et al. (1993), results showed that specific information on validity scales enabled many participants to avoid excessively elevating scores on both validity and clinical scales whereas information on CHI symptoms only served to inflate some of the clinical scale scores as well as the F scale score (Lamb et al., 1994). Participants receiving both types of information responded similarly to participants who received validity scale information only. The mean F scale score for experimental conditions varied, ranging from T scores of 65.5 (validity information only) to 89.5 (CHI information only), with the score for the group receiving both types of information (68.2) falling slightly above the score obtained by the group receiving validity scale information only. The mean Fb scale scores paralleled the scores reported for the F scale. Lamb et al. (1994) did not calculate effect sizes or compute optimum cutting scores for validity indicators.

Both Lamb et al. (1994) and Rogers et al. (1993) have demonstrated that coaching participants about detection strategies helps them to avoid detection at a substantially higher rate than participants provided only with information about a specific disorder.

In a meta-analysis of 15 MMPI-2 malingering studies, Rogers et al. (1994) examined the effect sizes obtained when fakers were compared to normals and to psychiatric patients (See Table A.2). Similar to the meta-analytic findings of Berry et al. (1991) for MMPI malingering studies, Rogers et

al. (1994) found that effect sizes varied greatly. However, unlike the findings of Berry et al. (1991), Rogers et al. (1994) found that the variation was not caused by sample type, type of design (within groups vs. between groups), or type of malingering instructions.

Rogers et al. (1994) found that only three scales had very large effect sizes (>2.00), F, F-K, and Obvious-Subtle subscales (O-S). The effect sizes reported for F and F-K converged with the findings of Berry et al. (1991), but the O-S findings were highly dissimilar. Berry et al. (1991) reported an overall effect size of 1.38, whereas Rogers et al. (1994) reported an overall effect size of 3.09. Although differences in O-S between these meta-analyses are difficult to understand, Rogers et al. (1994) noted that "changes between the MMPI and MMPI-2 in their respective T-score transformations clearly modify each scale's contribution to O-S" (p. 234).

In addition, Fb and Ds-r2 showed very large effect sizes for MMPI-2 comparisons between fakers and normals, and also showed large effect sizes (>1.50) for comparisons between fakers and psychiatric patients (Rogers et al., 1994). Rogers et al. (1994) lamented the fact that most MMPI-2 malingering studies have chosen to examine the 34-item version of the MMPI's Dsr scale (Ds-r2). Rogers et al. (1994) suggested "refining" the Ds-r2 scale, for example, by combining the remaining 58 items of the Ds scale with the 34 items of the Ds-r2 scale with the hope that it might produce a scale with high discriminability. The weakest malingering indicator identified in the meta-analysis by Rogers et al. (1994) was the Lachar-Wrobel critical items.

Rationale for the Present Study

The present study examined the effects of symptom information and validity scale information on the malingering of depression on the MMPI-2. Only two studies

have looked at the effect of specific symptom information combined with information about MMPI-2 validity scales on successful malingering on the MMPI-2 (Lamb et al., 1994; Rogers et al., 1993). However, neither of these studies examined the effect of information about depression and information about the validity scales on malingering depression on the MMPI-2. In addition to being the first study to examine the effects of information about depression and information about validity scales on malingering performance, this dissertation attempted to improve upon the research methodology used in similar studies.

Studying the ability of persons to malingering depression on the MMPI-2 is a relevant topic because major depression is very common in SSI disability claims (Griffin et al., 1997; Reich, 1986) as well as in personal injury lawsuits (Silvain, 1992; as cited in Lees-Haley & Dunn, 1994). Many individuals may recognize the advantages in either exaggerating existing sub-clinical depressive symptoms or in completely fabricating depression. Clearly, there are monetary and other incentives for a person to appear depressed, particularly in forensic settings and in situations where individuals are seeking disability status. A recent study examined dissimulation among SSI disability claimants in Los Angeles (Griffin et al., 1997). Thirty-six percent of 100 claimants tested complained of depression, 43% of memory impairment, 34% of poor concentration, and 34% of auditory or visual hallucinations. In the Griffin et al. (1997) study, malingering was assessed using a Composite Disability Malingering Index that was specifically created for the disability applicant population in Los Angeles. Specifically, the instrument took into account that many applicants exhibited borderline intellectual functioning and that they had the complaints listed above. The Composite Disability Malingering Index was composed of a combination of psychometric indices of malingering, including the MMPI

(11 items from the F scale), the M Test (10 items; a measure used to detect the malingering of schizophrenia), and the MCMI-II (4 items that reflected implausible yet "nonbizarre statements" that are endorsed by less than 1% in clinical populations). A cutting score was established that effectively differentiated between disabled nonmalingerers, instructed malingerers, and a population of disability applicants, some of whom were likely to be malingering and some of whom were not (Griffin et al., 1997). This cutting score classified 19% of the disability claimants to be malingering. It was judged that a 19% base rate of malingering for disability applicants was probably a high estimation for other regions of the country because of the particular circumstances in Los Angeles County operations regarding SSI disability applications (Griffin et al., 1997).¹² Clearly, malingering is an issue among SSI disability applicants and depression is one of the more frequent complaints of this population.

With the MMPI-2 being one of the most commonly used personality assessment measures, its effectiveness in the detection of malingering of all disorders must be thoroughly examined. Depression is one disorder that has not been studied much in terms of malingering on the MMPI-2 whereas studies looking at PTSD and schizophrenia are more common.

Because it is unclear what kinds of information malingerers obtain or how they prepare to mangle, studying the effect of providing depression symptom information and validity scale information on malingering performance is an important avenue of research. Clearly, not all malingerers

¹²Griffin et al. (1997) indicated that because of county inconsistency and inefficiency, claimants have "sometimes tacitly and sometimes explicitly indicated their conviction that the program is such that one must exaggerate or lie in order to get disability funds in Los Angeles" (p. 1426).

are sophisticated or motivated enough to gather the types of extensive information that will be provided to participants in this experiment. But, by allowing participants to study validity scale and symptom information, they may be as well-prepared as the most sophisticated and motivated real-life malingerer. Taking into account that the participants in this study will probably be better prepared to fake on the MMPI-2 than the average malingerer, this study will be a stringent test of the MMPI-2's ability to detect malingering.

Although current studies suggest that providing symptom information does not aid persons in malingering successfully on the MMPI-2, no studies have systematically examined the effect of symptom information on malingering of major depression. In a study focusing on faking borderline personality disorder on the MMPI-2, Sivec et al. (1995) included a malingering depression group and provided them with information on depression symptoms. Since BPD has been conceptualized as a subaffective disorder, the depression group was used in that study as a way of looking at the differences between simulated affective disorder and simulated BPD (i.e., a comparison group). There were no other experimental conditions involving malingering depression in that study.

An unpublished study by Petersen and Viglione (1991) cited in Rogers et al. (1993), "looked at the effects of psychological training (clinical psychology graduate students versus non-mental health professionals) on the ability to fake specific disorders on the MMPI" (p. 216). In his brief discussion of that study, Rogers et al. (1993) did not mention if symptom information was provided to participants. According to Rogers et al. (1993), results suggested that depression and other non-psychotic disorders are easier to fabricate on the MMPI than psychotic spectrum disorders. These results may imply that the public is

better informed about depressive symptomatology than about psychotic disorders. Indeed, many untrained individuals can accurately endorse enough symptoms of depression to meet DSM-III-R criteria for Major Depression (Lees-Haley & Dunn, 1994). Depression is a fairly common disorder in relation to other disorders, with lifetime prevalence rates in community samples ranging from 10%-25% for women and 5%-12% for men (DSM-IV, 1994). Therefore, most people either have personal experience with depression or know a family member or friend who has had depression. Having this knowledge may be beneficial in faking depression on the MMPI-2.

In designing this study, an effort was made to improve upon the research methodology of previous studies. First, this study employed a scenario that is more appropriate and relevant to the college population than previous studies. Those studies have used scenarios asking students to imagine themselves in situations more suitable to a community sample. For example, Lamb et al. (1994) asked students to imagine that they had filed a lawsuit after suffering a mild head injury in a car accident caused by a drunk truck driver. They were asked to imagine further that, after being told that they would have to be psychologically evaluated (including the MMPI-2) as part of the lawsuit, "their attorney had informed them that the greater the trauma attributable to the accident, the larger the potential financial settlement" (Lamb et al., 1994, p. 9). To maximize subject involvement and motivation, the present scenario was written with college students specifically in mind. Participants were asked to imagine feeling slightly depressed due to increasingly bad grades, fear of losing their academic scholarship, and financial difficulties. Participants were requested to imagine further that they discovered that they might be eligible to receive monthly payments as well as deadline extensions and rescheduling of

exams from professors if they were found to be severely depressed in a psychological evaluation including the MMPI-2. In contrast to some previous research which used relatively brief and uninvolved scenarios (Rogers et al., 1993; Wetter et al., 1993; Wetter et al., 1994), care was taken to provide a longer scenario to enable participants to more easily imagine themselves in the situation suggested in the scenario.

Another difference between the current study and previous malingering research pertains to the nature of the symptom information provided to participants. Although many studies have given participants symptom information based on DSM-III-R (1987) or DSM-IV (1994) criteria, few of them have mentioned the fact that persons do not have to have all the symptoms listed in order to receive the diagnosis. For example, in the study by Sivec et al. (1995) the depression information was translated from the DSM-III-R (1987) criteria for major depression into a narrative in the second person:

You feel very sad and depressed nearly every day.
You do not feel interested in most things and find
it difficult to find anything pleasurable. You have
recently noticed a significant change in weight.
(p. 300)

Unfortunately, these types of symptom descriptions may unwittingly lead a person to exaggerate the disorder on the MMPI-2, thus leading to an inflated malingering detection rate. The current study presented the numbered DSM-IV criteria in a simplified and easy to understand manner, and clearly delineated the number of symptoms needed for a diagnosis of major depression. It may be argued that persons preparing to mangle may study lay literature on the specific disorder (or may not consult outside sources at all), and thus, would not have knowledge of the specific criteria used in the DSM-IV. However, the sophisticated malingerer would have little difficulty obtaining access to

psychology textbooks and the DSM-IV. These sources spell out the criteria and clearly state that only a certain number of the criteria are necessary for diagnosis. Providing complete and understandable information about depression to participants--material which is readily available to the sophisticated malingerer--may make participants better able to fake depression on the MMPI-2 than many of their real-life counterparts. Assuming that symptom information helps malingerers fake more successfully and that a minority of malingerers study the disorder they are attempting to fake, supplying participants with symptom information is a stringent test of the ability of the MMPI-2 to detect malingering.

This study also sought to improve upon past research by providing more understandable validity scale information, with an emphasis on strategies for eluding detection. In their information about validity scales and malingering strategies, Lamb et al. (1994) and Rogers et al. (1993) discussed validity scale names and, in some cases, how the validity scales were developed. The validity information used in this study eliminated extraneous information (e.g., scale names and scale development and history) and the use of technical terms such as "malingering." The focus of the information was on strategies for successful malingering. Although this type of simplified MMPI-2 validity information may not be reflective of kind of information that is readily available to the typical malingerer in the real world, these detection strategies may be learned by the more intelligent and sophisticated malingerer through studying a detailed text on the MMPI-2. Assuming that validity scale information helps malingerers fake more successfully and that most malingerers do not have knowledge of the MMPI-2 validity scales, this study, in providing validity scale information to faking participants, was a rigorous test of the MMPI-2's efficacy in detecting malingering11.

Measures were also taken in this dissertation research to make sure that participants truly understood the strategies for eluding detection. Similar to the Lamb et al. (1994) study, the present study provided time for the participant to absorb the information presented to them and had them process the information by filling out a worksheet on the information. In addition, participants were quizzed on the information to test their comprehension of the material. The quiz focused solely on the strategies for successful malingering and, unlike Lamb et al. (1994), did not include questions about scale development or scale names. Therefore, persons who passed the quiz ($\geq 80\%$ correct) were likely to have had a firm grasp of how to elude detection.

In line with Rogers' (1988) recommendations, a further improvement on existing methodology in this study involved employing a more extensive feedback questionnaire. The questionnaire was based partially on the debriefing materials used by Lamb et al. (1994), with questions that evaluated perceived ability to stay engaged in the task and perceptions of how well they were able to stay in the role. In addition, the feedback questionnaire used in this dissertation assessed strategies used in attempting to successfully malingering using open-ended questions and checklists. In particular, the feedback questionnaire strategy questions were created to learn what strategies were independently employed to "fake out" the MMPI-2 by successful participants who did not receive depression information or validity information (suggested by Sivec et al., 1994).

To date, no feedback questionnaires used in MMPI-2 malingering research have investigated participants' personal experience with mental disorders. Therefore, in an exploratory effort to assess the effect of personal experience with depression on malingering success,

participants were asked if family members, friends, or they themselves had actually had depression. In addition, they were asked the duration of the depressive episode, when it occurred, and if psychotropic medications were prescribed to combat the depression. Although severity or existence of clinical level depression cannot be inferred from such self-report information, relationships between MMPI-2 profiles and this information were examined to see if more systematic study of this area is warranted. For example, if participants who reported being depressed in the past are shown to be more successful at malingering depression on the MMPI-2 than participants who did not report previous episodes of depression, future research may want to investigate the effects of a past experience of a disorder on the successful malingering of that disorder at a later date.

A major limitation of MMPI and MMPI-2 research on malingering is the use of college students as participants in both experimental and control conditions and the lack of clinical comparison groups. Although most of the participants in this study were college students, one group was composed of clinically depressed inpatients. The use of a clinical comparison group allowed the current study to see if fakers could be differentiated not only from normal participants, but also from psychiatric patients.

Another limitation of some previous research in this area is small sample sizes (see Rogers et al., 1993). By using mostly college students as study participants, a large sample was easily obtained thus increasing statistical power. Also, most studies on malingering on the MMPI-2 have mixed gender groups. Including a mix of men and women in the same groups dilute findings since it is well known the women and men manifest depression in different ways and because women and men respond differently to the MMPI-2. In this study, there were enough male and female participants

to analyze the data for men and women separately.

Providing participants with MMPI-2 validity scale information is a potential ethical problem. Ben-Porath (1994) and Berry, Lamb, Wetter, Baer, and Widiger (1994) have described the ethical issues pertaining to malingering research that involves coaching participants on detection strategies. The central dilemma outlined by Ben-Porath (1994) involves the conflict between the need for such research and the possibility that published journal results might aid potential malingerers in successful malingering. Although there are no easy solutions to this predicament, Berry et al. (1994) suggested limiting the amount of information regarding detection strategies in published studies. Although this would make it difficult for researchers or clinicians to evaluate the studies providing malingering detection information, it would keep detection strategies from appearing in readily accessible journals, thus safeguarding the integrity of the test. Interested researchers who are ethically bound by the APA ethical code could obtain detailed information by writing directly to the authors.¹³

¹³ Because of these ethical issues, the information pertaining to the validity information in this study were not printed in this dissertation. Specifically, the validity information sheets, the validity information worksheets, the validity worksheet answers, and the validity quizzes were omitted from the document. Also, detailed descriptions of the validity information were omitted. Interested researchers who are ethically bound by the APA ethical code can write the author or Jim Clopton, Ph.D. at Texas Tech University to receive more details about the validity information provided to some of the participants in this study.

Table A.1

Demographic and Methodological Characteristics and Mean D for 24 MMPI Studies of Malingering

Citation	Group	N	% Male	Instruction	Design	Comparison	Mean <u>D</u>
Anthony (1971)	B	40	100	1	RM	3	1.3
	M	32	100	1	BG	3	0.3
	C	32	100	3			
Ashlock (1987)	M	5	100	5	BG	2	1.8
	C	21	100	3			
Burkhart (1978)	B	60	50	1	RM	1	0.4
Butler (1969)	B	8	0	2	RM	1	1.2
Cofer (1949)	B	28	-	1	RM	1	5.5
	M	28	-	1	BG	1	5.6
	C	26	-	3			
Exner (1963)	B	25	45	1	RM	1	2.5
	M	25	45	1	BG	1	3.8
	C	25	45	3			
Fairbank (1985)	M	15	100	2	BG	2	3.3
	C	15	100	3			
	M	15	100	2	BG	2	2.4
	C	15	100	3			

Table A.1

Continued.

Citation	Group	N	% Male	Instruction	Design	Comparison	Mean $\bar{D}$
Gallucci (1984)	B	84	100	4	BG	4	0.6
Gendreau (1973)	B	24	100	1	RM	1	2.4
Gough (1947)	B	11	100	1	RM	1	3.4
	B	11	100	1	RM	1	8.1
	M	11	100	1	BG	2	0.8
	C	57	100	3			
	M	11	100	1	BG	2	1.4
	C	13	100	3			
Gough (1950)	M	319	-	1	BG	1	2.9
	C	960	-	3			
	M	319	-	1	BG	2	2.3
	C	500	-	3			
	M	319	-	1	BG	2	1.7
	C	313	-	3			
Gough (1954)	M	354	-	1	BG	2	3.7
	C	915	-	3			
	M	354	-	1	BG	1	4.4
	C	507	-	3			

Table A.1

Continued.

Citation	Group	N	% Male	Instruction	Design	Comparison	Mean $\bar{D}$
Grossman (1988)	M C	49 52	80 80	4 3	BG	4	0.5
Grow (1980)	M C M C M C	50 50 16 14	- - - -	1 3 5 3	BG BG BG	1 2	3.1 2.2
Hawk (1989)	M C M C	18 17 18 36	- - - -	5 3 5 3	BG BG BG	2 1	0.7 0.7
Heaton (1978)	M C	16 16	70 80	2 3	BG	2	1.0
Hsu (1989)	M C	40 20	56 56	1 3	BG	1	1.5
McCaffrey (1989)	M C M C	12 11 12 11	100 100 100 100	2 3 2 3	BG BG	2 2	2.4 1.9
Peterson (1989)	B	20	30	1	RM	1	0.8

Table A.1

Continued.

Citation	Group	N	% Male	Instruction	Design	Comparison	Mean <u>D</u>
Pollack (1984)	M	26	50	4	BG	4	1.6
	C	26	50	3			
	M	26	50	4	BG	4	0.7
	C	26	50	3			
Rice (1983)	B	25	100	1	RM	3	2.8
Walters (1988a)	M	35	100	4			
	C	30	100	3	BG	2	1.0
Walters (1988b)	M	12	100	5			
	C	20	100	3	BG	4	0.6
Waslyiw (1988)	M	35	100	4			
	C	39	100	3	BG	4	0.6

Note. - = not reported; Group M = malingering. C = control (honest). B = both (e.g., both malingered and honest in a repeated measures design); Instruction (for repeated measures design, only malingering instruction type coded) 1 = malingering global mental illness, 2 = malingering specific mental illness syndrome, 3 = standard (honest), 4 = standard but presumed to be malingering because of circumstances, 5 = standard but judged malingering using external criterion; Design RM = repeated measures, BG = between group; Comparison 1 = Normals vs. Normals malingering, 2 = Normals malingering vs. Patients honest, 3 = Patients honest vs. Patients exaggerating, 4 = Inferred response set, indeterminate group membership.

Table A.1

Continued.

Note. From "Detection of malingering on the MMPI: A meta-analysis" by D.T.R. Berry, R.S. Baer, and M.J. Harris, 1991, Clinical Psychology Review, 11, p. 590 - 591. Adapted by permission.

Table A.2

Demographic and Methodological Characteristics and Mean D for 20 MMPI Studies of Malingering

Citation	Group	N	% Male	Instruction	Comparison	Mean D
Austin (1992)	S	37	-	2	1	2.61
	C	33	-	4		
Bagby, Rogers, Buis & Kalemba (1994)	S	58	31	2	1	1.72
	C	90	31	4	2	1.54
	P	95	52	4		
Bagby, Rogers, & Buis (1994)	S	31	34	2	1	2.41
	C	47	34	4	2	1.64
	P	66	76	4		
Bagby, Buis, & Nicholson (1995)	S	74	39	2	1	-
	P	128	50	4	2	-
	P	157	82	4	2	-
	C	198	30	4		
Cassissi & Workman (1992)	S	20	58	2	1	1.66
	C	20	58	4		
Graham, Watts, & Timbrook (1991)	S	50	54	2	1	2.28
	C	50	54	4	2	1.33
	P	50	60	4		

Table A.2

Continued.

Citation	Group	N	% Male	Instruction	Comparison	Mean $\bar{D}$
Iverson, Franzen & Hammond (1995)	S	53	100	2	1	-
	P	51	100	4	2	-
	C	53	100	4		-
Lamb, Berry, Wetter, & Baer (1994)	S	159	47	3	1	-
	C	86	53	4		
Lees-Haley (1991)	M	26	50	1	3	2.12
	C	21	33	4		
Lees-Haley (1992)	M	55	58	1	3	2.52
	C	64	42	4		
Lindblad (1993)	P	66	100	4	4	3.33
				2		
Rogers, Bagby, & Chakraborty (1993)	S	59	49	2	1	1.82
	C	13	49	4	2	1.60
	P	37	84	4		
Rogers, Sewell, & Ustad (1993)	P	42	49	4	2	1.29
				2		
Sivec, Lynn, & Garske (1993)	S	179	40	3	1	2.08
	C	58	40	4		

Table A.2

Continued.

Citation	Group	N	% Male	Instruction	Comparison	Mean $\bar{D}$
Sivec, Hilsenroth, & Lynn (1995)	S S P C	65 62 40 61	25 16 13 16	3 3 4 4	1 2 2	- - -
Timbrook, Graham, Keiller, & Watts (1993)	S P	47 47	57 57	2 2	5 2	4.53
164 Wetter, Baer, Berry, Smith, & Larsen (1992)	S C	70 68	43 49	5	1	2.68
Wetter, Baer, Berry, Robinson, & Sumpter (1993) (s1) ^a	S P	20 20	40 70	3 4	2	1.08
Wetter, Baer, Berry, Robinson & Sumpter (1993) (s2) ^a	S P	22 20	68 55	3 4	2	2.85

Table A.2

Continued.

Citation	Group	N	% Male	Instruction	Comparison	Mean <u>D</u>
Wetter, Baer, Berry, & Reynolds (1994)	S	23	13	3	1	-
	S	23	22	2	2	-
	P	36	17	4	2	-
	C	36	31	4	2	-

Note. - = not reported; Groups: M = malingering. C = control. P = psychiatric. F = forensic (inmates); Instructions: 1 = standard instructions (for suspected malingerers), 2 = malingering global mental illness, 3 = malingering specific mental illness syndrome, 4 = standard instructions (for honest test-takers), 5 = malingering either moderate or severe mental illness (combined); Comparison type: 1 = Simulators vs. Normals, 2 = Simulators vs. Patients (honest), 3 = Malingerers vs. Normals, 4 = Repeated measures (same clinical sample with two administrations under honest and fake instructions), 5 = Repeated measures (same student sample with two administrations under honest and fake instructions).
Note. From "A meta-analysis of malingering on the MMPI-2" by R. Rogers, 1994, Assessment, 1, p. 227 - 237. Currently seeking permission to adapt table.

APPENDIX B
INFORMED CONSENT

APPENDIX B.1: INFORMED CONSENT FORM
FOR FAKING GROUPS

INFORMED CONSENT (A)

MMPI-2 Study

I freely and voluntarily consent to be a participant in the research project entitled "The Effect of Symptom Information and Validity Scale Information on the Malingered Depression on the MMPI-2" to be conducted at Texas Tech University during the 1996 Fall semester with Gina Walters, M.A. as the principal investigator. I have been told that this experiment will last approximately 3 hours.

I understand that the purpose of this research is to study the way in which certain types of information may affect the way people respond to a psychological test (the MMPI-2).

I understand the research procedures will be as follows: I will be asked to read a scenario and then to role-play as if I actually had that experience as I fill out the MMPI-2. I may also be asked to learn some additional information and use it while I take the MMPI-2. After completing the MMPI-2, I will be asked to answer a brief questionnaire.

I also understand that cash prizes will be awarded to the five best role players who participate in this experiment. The prizes are: 1st prize is \$75, 2nd - \$30, 3rd - \$20, 4th - \$15, and 5th - \$10. I have been told my individual responses will be kept strictly confidential. All scores will be identified only by a code number, and my individual performance will not be revealed to anyone without my expressed permission. I understand that the principal investigator, Gina Walters, will have a master list identifying participants by code number for the purpose of awarding the prizes, but once the prizes have been received, the master list will be destroyed.

I understand that I may withdraw my consent and discontinue participation in this research at any time without receiving a "no-show" or any other negative consequences. I will receive experimental credit upon completion of the session. I have been given the opportunity to ask questions concerning the procedures for the study, and any questions have been answered to my satisfaction.

I understand that if I desire further information about these procedures or this experiment, I should contact Gina Walters in the psychology department by leaving a note in her box. I understand that this research is being supervised by Dr. Jim Clopton who is also able to answer questions concerning these procedures. He can be reached by leaving a message at 742-3711.

I understand that the Texas Tech University Institutional Review Board (IRB) for the Protection of Human Participants has been established for the protection of individuals participating in research projects conducted by university personnel. I may contact the IRB by writing them in care of the Office of Research Services, Texas Tech University, Lubbock, Texas 79409, or by calling (806) 742-3884.

I understand that there are no known risks involved in my participation in this experiment. However, if any injury were to occur, treatment for the injury would not necessarily be available at Texas Tech University or the Student Health Center, nor is there necessarily any insurance to cover such an injury. Financial compensation for any such injury must be provided through the participant's own insurance program. Further information about these matters may be obtained from Dr. Robert M. Sweazy, Vice Provost for Research, 742-3884, Room 203 Holden Hall, Texas Tech University, Lubbock, Texas 79409-1035.

I have read and understand the above.

Participant's signature

Date

APPENDIX B.2: INFORMED CONSENT FORM
FOR HONEST RESPONDING GROUPS

INFORMED CONSENT (B)

MMPI-2 Study

I freely and voluntarily consent to be a participant in the research project entitled "The Effect of Symptom Information and Validity Scale Information on the Malingered Depression on the MMPI-2" to be conducted at Texas Tech University during the 1996 Fall semester with Gina Walters, M.A. as the principal investigator. I have been told that this experiment will last approximately 2 hours.

I understand that the purpose of this research is to study the way in which certain types of information may affect the way people respond to a psychological test (the MMPI-2).

I understand the research procedures will be as follows: I will be asked to fill out the MMPI-2 in an honest manner. After completing the MMPI-2, I will be asked to answer three brief questionnaires. I have been told my individual responses will be kept strictly confidential. All scores will be identified only by a code number, and my individual performance will not be revealed to anyone without my expressed permission.

I understand that I may withdraw my consent and discontinue participation in this research at any time without receiving a "no-show" or any other negative consequences. I will receive experimental credit upon completion of the session. I have been given the opportunity to ask questions concerning the procedures for the study, and any questions have been answered to my satisfaction.

I understand that if I desire further information about these procedures or this experiment, I should contact Gina Walters in the psychology department by leaving a note in her box. I understand that this research is being supervised by Dr. James Clopton who is also able to answer questions concerning these procedures. He can be reached by leaving a message at 742-3711.

I understand that the Texas Tech University Institutional Review Board (IRB) for the Protection of Human Participants has been established for the protection of individuals participating in research projects conducted by university personnel. I may contact the IRB by writing them

in care of the Office of Research Services, Texas Tech University, Lubbock, Texas 79409, or by calling (806) 742-3884.

I understand that there are no known risks involved in my participation in this experiment. However, if any injury were to occur, treatment for the injury would not necessarily be available at Texas Tech University or the Student Health Center, nor is there necessarily any insurance to cover such an injury. Financial compensation for any such injury must be provided through the participant's own insurance program. Further information about these matters may be obtained from Dr. Robert M. Sweazy, Vice Provost for Research, 742-3884, Room 203 Holden Hall, Texas Tech University, Lubbock, Texas 79409-1035.

I have read and understand the above.

Participant's signature

Date

APPENDIX C

MATERIALS USED IN PARTICIPANT GROUPS

APPENDIX C.1: DEMOGRAPHIC INFORMATION SHEET

DEMOGRAPHIC INFORMATION SHEET

Age: _____

Year in school:

Freshman	_____
Sophomore	_____
Junior	_____
Senior	_____

Race (check all that apply):

Asian	_____
African-American	_____
Hispanic	_____
Native American	_____
White	_____
Other	_____

APPENDIX C.2: EXPERIMENTAL SCENARIO

SCENARIO

Imagine that you have been feeling a bit down lately because you are not doing as well as you would like in your classes.

You just got back from class where you got a test back . . . ANOTHER C. You are used to making B's and some A's, but lately, all you can seem to make are C's. Your grades are not as good this year because you are taking some difficult classes. To make matters worse, you are in danger of losing your scholarship money as a result of your declining grades. The loss of this financial aid could force you to drop out of college. And, on top of this, you are ALREADY finding yourself short of money to pay the bills, which is causing your parents to be irritated with you. Needless to say, you are feeling much pressure to bring up your grades. If you just had some extra time to catch up on your reading and more time to study for tests, you feel confident you could raise your grades to B's this semester.

When you are telling some friends about your situation and how badly you feel about it, they tell you that they know a student who was depressed, and, as a result, was able to get a break from professors. For example, this student was given extra time to finish papers and was able to postpone exams. In addition, the student was able to receive extra money each month from a local agency because of the depression. After checking around, you discover that if you are tested by a psychologist and your testing indicates that you are severely depressed, the testing center will write official letters to your professors that require them to give you special accommodations in meeting the requirements of the course (such as extra time completing a paper or letting you reschedule an exam). Also, you might be eligible to receive a monthly disability check of \$200 for at least several months from a local agency. Obviously, this would be helpful to you in paying off those bills and would enable you to have time to bring up your grades in your classes, thus helping you to keep your scholarship. You will be able to stay in school after all if this all works out!

You decide to go over to complete the depression testing. The testing center tells you that you must take a battery of personality tests, including the MMPI-2, to assess the severity of your depression. You learn that the testing must indicate that you are severely depressed in order for you to receive the financial and academic assistance that you have heard about. However, you are not feeling THAT depressed - you just feel a bit down several times a week.

But, you decide to go ahead with the testing anyway and have decided to exaggerate your depressive symptoms so that you can get some short-term help with your situation. After all, you have a lot on the line here - you are trying to stay in school and hoping to be able to catch up on your bills.

Pretend that you are now at the testing center preparing to take the MMPI-2. Your goal is to answer the MMPI-2 questions in such a way as to make your minor depressive symptoms seem much worse than they actually are. Keep in mind that it is important to convince the person interpreting the test that you are **severely depressed** so that you can receive financial assistance and help from your professors. However, it will be obvious if you over-exaggerate, so it is very important that you answer the MMPI-2 in a **believable** manner.

As you can see, we want you to imagine that you are the student I just described to you. We want you to take the role of a person who is trying to fake depression on the MMPI-2.

****DO **NOT** turn the page until told to do so****

APPENDIX C.3: DEPRESSION INFORMATION

SYMPTOMS OF DEPRESSION

In order to be considered to have depression, you MUST have ONE or BOTH of the following 2 symptoms:

1) You have a depressed mood most of the day, nearly every day for at least the last 2 weeks. For example, you might feel sad or empty or cry a lot nearly every day.

2) You are not interested in things and you can't think of anything that would be fun or make you feel good. For at least the last 2 weeks you have felt this way. Almost all of the activities you usually enjoyed do not seem pleasurable to you now.

To be considered to have depression, you must also have had FOUR or more of the following symptoms during the last 2 weeks:

3) A significant change in your weight or appetite. An increase or decrease in your normal appetite must be noticed nearly every day for the last 2 weeks. If you lost weight, it cannot be a result of dieting.

4) Your sleep has been disturbed nearly every night for the last 2 weeks. Your sleep disturbance fits one of two opposite patterns: either you have had trouble sleeping (insomnia), or you have been sleeping too much.

5) Your physical activity level has been noticeably different from normal every day for the last 2 weeks. That change in your activity level fits one of these two opposite patterns: either you have felt "slowed down" or restless and fidgety.

6) You feel very tired or worn out nearly every day for the last 2 weeks.

7) You feel like you are worthless or feel very guilty nearly every day for the last 2 weeks.

8) You find it difficult to concentrate and to think. You have had difficulty making decisions every day for the last 2 weeks.

9) You think about death and suicide often.

In addition, these symptoms must "get in the way" with or interfere with your life. For example, you might find it hard to get your school work done or you might find that

your relationships are suffering because of your depression. These symptoms are also very upsetting to you and represent a change in your normal functioning. Remember, someone has depression if they have one or both of symptom #1 and #2, AND also have at least 4 of the other symptoms (#3 - #9) listed above.

APPENDIX C.4: DEPRESSION INFORMATION WORKSHEET

DEPRESSION WORKSHEET

1. List as many symptoms of depression as you can - there are nine in all.

2. A person who is diagnosed with Depression **must** have one of two symptoms. Please describe those 2 symptoms.

3. Please describe two different ways that depression can affect sleeping patterns.

4. Please describe two different ways that depression can affect appetite.

5. Please describe two different ways that depression can affect energy level.

APPENDIX C.5: DEPRESSION WORKSHEET ANSWERS

DEPRESSION WORKSHEET ANSWERS

1. Symptoms of depression:

- 1) **Depressed mood** most of the day, nearly every day.
- 2) **Lack of interest** in activities that you usually enjoy.
- 3) Significant **change in your weight or appetite**. An increase or decrease in appetite must be noticed nearly every day for the last two weeks.
- 4) **Sleep problems**: either trouble sleeping nearly every night or sleeping too much every day.
- 5) **Physically "slowed down"** every day for the last 2 weeks. Or, the opposite may be true: You feel **physically restless** and fidgety
- 6) Feeling very **tired** or worn out nearly every day.
- 7) Feeling **worthless or guilty**
- 8) **Difficulty concentrating** and making decisions
- 9) **Suicidal thoughts**

2. A person diagnosed with Depression must have one of the following 2 symptoms:

- 1) **Depressed mood** most of the day, nearly every day for at least the last 2 weeks.
- 2) **Lack of interest** in participating in activities that you usually enjoy. For at least the last 2 weeks, you are having a hard time thinking of anything that would be fun or make you feel good.

3. Depression can affect sleeping patterns in one of two ways:

Either you have **trouble sleeping** nearly every night or you may find that you have been **sleeping too much** every day. One of these two patterns has been apparent for at least 2 weeks.

4. Depression can affect appetite in one of two ways:

An **increase or decrease in appetite** must be noticed nearly every day for the last 2 weeks.

5. Depression can affect your energy level in one of two ways:

Feeling **physically "slowed down"** every day for the last 2 weeks. Or, the opposite may be true: Feeling **physically restless and fidgety** every day for the last 2 weeks.

APPENDIX C.6: DEPRESSION QUIZ

DEPRESSION QUIZ

1. It is common for a person with depression to feel guilty.

- A) True B) False

2. Of the following, which is not a common symptom of depression?

- A) Feeling guilty
B) Crying every day
C) Trouble concentrating
D) Feeling angry

3. Even if your depressive symptoms do not interfere with your life and do not cause you much emotional distress, you can still receive the depression diagnosis.

- A) True B) False

4. A person with depression frequently feels like he or she can't get going and may actually move in a slowed down manner.

- A) True B) False

5. You must have all nine symptoms of depression to be formally diagnosed with depression.

- A) True B) False

6. Which of the following is NOT a common physical symptom of depression?

- A) Having no energy
B) Frequent headaches
C) Moving slowly
D) Significant change in appetite

7. A depressed person may have difficulty concentrating.

- A) True B) False

8. Most of the time, depressed persons can think of something to do that would be fun or would make themselves feel good.

A) True

B) False

9. A person with depression may have trouble sleeping.

A) True

B) False

10. Persons with depression:

A) Are often upset about their depressive symptoms

B) Often find it hard to think

C) Often feel tired or worn out

D) All of the above

APPENDIX C.7: FEEDBACK QUESTIONNAIRE FOR FAKING GROUPS¹⁴

VAL/DEP

FEEDBACK QUESTIONNAIRE

Please circle the appropriate number:

1) How well did you understand the instructions (i.e. did you understand what you were asked to do?)

1	2	3	4	5
completely understood				Did not understand at all

2) How motivated were you to try to appear depressed on the MMPI-2?

1	2	3	4	5
Did not Try				Very Motivated

3) How much of the test did you answer following the instructions you were given (i.e., to take on the role of a depressed person)?

1	2	3	4	5
None 0%		Half of the test 50%		All of test 100%

4) How successful do you think you were in appearing depressed on this test?

1	2	3	4	5
not successful at all				totally successful

¹⁴ There were four different versions of the feedback questionnaire. The version of the feedback questionnaire that is provided in this appendix is the one that the validity and depression information group received because it is the most extensive. The three other faking groups did not have question 7b. In place of 7b, the validity information group and the no information group were asked an open ended question ("How did you know which items to answer on the MMPI-2 in order to appear depressed? In other words, where did you learn about depression?"). The honest responding group questionnaire did not have questions 2-5, 7a, or 7b.

5) How did you attempt to appear depressed? In other words, what strategies did you use to fake out the MMPI-2? Please answer this question before turning to the next page.

6) Many people who complete a long questionnaire like the MMPI-2 find that they have difficulty concentrating throughout the entire test and answer some or all of the questions without paying much attention. We would like you to tell us how much of a problem this was for you while you were completing the test today. Please be as honest as possible. Remember that your responses on the test and this sheet are anonymous. There is no penalty for admitting that you became bored or stopped working carefully at some point. We just want to know how much of a problem this was for you. Your honest answer will help us understand your results.

Please indicate the percentage of questions that you answered without paying attention to the question: _____

7a) Please check ALL the ways in which you attempted to fake being depressed on the MMPI-2 :

- _____ marked ALL items that sounded like depression
- _____ marked only SOME items that sounded like depression
- _____ marked items that any person with a severe mental illness would endorse (ie. hearing voices, etc.)
- _____ made a conscious attempt to not over-exaggerate severity of depression

7b) Mark the statement which best describes your strategy for faking depression on the MMPI-2.

- _____ used information about the validity scales on the MMPI-2
- _____ used information about depression symptoms
- _____ used both validity scale information and depression symptom information about equally

***Please read the following before proceeding with this questionnaire:

SYMPTOMS OF DEPRESSION

In order to be considered to have depression, you MUST have ONE or BOTH of the following 2 symptoms:

1) You have a depressed mood most of the day, nearly every day for at least the last 2 weeks. For example, you might feel sad or empty or cry a lot nearly every day.

2) You are not interested in things and you can't think of anything that would be fun or make you feel good. For at least the last 2 weeks you have felt this way. Almost all of the activities you usually enjoyed do not seem pleasurable to you now.

To be considered to have depression, you must also have had FOUR or more of the following symptoms during the last 2 weeks:

3) A significant change in your weight or appetite. An increase or decrease in your normal appetite must be noticed nearly every day for the last 2 weeks. If you lost weight, it cannot be a result of dieting.

4) Your sleep has been disturbed nearly every night for the last 2 weeks. Your sleep disturbance fits one of two opposite patterns: either you have had trouble sleeping (insomnia), or you have been sleeping too much.

5) Your physical activity level has been noticeably different from normal every day for the last 2 weeks. That change in your activity level fits one of these two opposite patterns: either you have felt "slowed down" or restless and fidgety.

6) You feel very tired or worn out nearly every day for the last 2 weeks.

7) You feel like you are worthless or feel very guilty nearly every day for the last 2 weeks.

8) You find it difficult to concentrate and to think. You have had difficulty making decisions every day for the last 2 weeks.

9) You think about death and suicide often.

In addition, these symptoms must "get in the way" with or interfere with your life. For example, you might find it hard to get your school work done or you might find that

your relationships are suffering because of your depression. These symptoms are also very upsetting to you and represent a change in your normal functioning. Remember, someone has depression if they have one or both of symptom #1 and #2, AND also have at least 4 of the other symptoms (#3 - #9) listed above.

9) Have you ever had depression?

YES, in the past _____ YES, currently _____ NO _____

IF YES:

How long did you have depression?

Were you diagnosed with depression? _____

How old were you when you had depression? (range of ages acceptable) _____

10) Have you ever been close to someone who had depression?

YES _____ NO _____

IF YES:

How long did the person have depression?

What is this person's relationship to you? _____

How old were you when a person close to you had depression? (range of ages acceptable) _____

11) Have you had any past or current psychiatric or psychological treatment? (ie. therapy, counseling, or hospitalization)

YES _____ NO _____

12) Have you ever taken anti-depressants (i.e., Zoloft, Prozac, etc.) or been treated for depression?

YES _____ NO _____

If you answered "yes" to the last question, briefly describe the treatment you received.

13) Have you ever taken the MMPI-2 before today?

YES _____ NO _____ UNCERTAIN _____

If Yes, how long ago _____

Please write comments or observations about participating in this experiment below.

*******REMEMBER, PLEASE DO NOT DISCUSS THIS EXPERIMENT WITH OTHER PSYCHOLOGY 1300 STUDENTS UNTIL THE END OF THIS SEMESTER. THIS IS VERY IMPORTANT IF THE RESULTS ARE TO BE INTERPRETABLE AND VALID.*******

APPENDIX C.8: DEBRIFING SHEET

DEBRIEFING SHEET

The MMPI-2 is one of the most widely used psychological tests. There are many situations in which a person taking the MMPI-2 may try to fake a mental disorder: criminals trying to appear insane, persons trying to obtain disability payments from the government, and persons pursuing personal injury lawsuits. Sometimes persons taking the MMPI-2 may locate information that can help them to fake a mental disorder more successfully. For example, a person wishing to fake depression on the MMPI-2 might read a book on the disorder or a book on the MMPI-2 itself.

This study was designed to find out if different types of information help people to better fake depression, and if this information affects the MMPI-2's ability to detect such exaggeration. To answer these questions, two types of information were provided: one group was told about the ways the MMPI-2 tries to detect when people are exaggerating, another group learned about the different symptoms associated with depression, a third group received both types of information, and a fourth group did not get either type of information. The four groups of participants were asked to fake depression on the MMPI-2. A fifth group of participants were asked to answer the MMPI-2 in an honest manner. We predict that the information will help people to better simulate being depressed, but that most people will still be detected by the MMPI-2. This research may help psychologists detect when someone is exaggerating depression in order to gain money in court or obtain disability payments. Perhaps more importantly, it may allow psychologists to have greater confidence in the test results and thus present a more convincing case for the client who is really suffering from depression.

There is a possibility that your participation in this experiment could compromise the validity of the test results if you later find it necessary to take the MMPI-2. However, if you do find it necessary to take the MMPI-2 or any other psychological test in the future, it is in your best interest to respond honestly. Fake responses, such as those suggested during this experiment, are usually detected and could jeopardize your results. Texas Tech University and the Department of Psychology do not condone attempts to deliberately manipulate test results.

Thank you for your help in this study. It would not be possible to continue psychological research without your cooperation. We hope you enjoyed this experiment and found it interesting to participate. If you would like to learn

more about this topic, you may contact me. We expect to have the results analyzed by the 1997 spring semester, so if you are curious about what we found, please leave me a message with your name and address in my box (labeled "WALTERS") on the first floor of the psychology building (near the coke machine).

Counseling opportunities: If you feel that you or someone you know has depression or some other problem that could be helped with therapy, there are several places to call for information. The counseling center located in West Hall (742-3674) provides free services for students. The Psychology Clinic (742-3737) in the psychology building provides services for a small fee.

APPENDIX D
VERBAL INSTRUCTIONS

APPENDIX D.1: VERBAL INSTRUCTIONS
FOR FAKING GROUPS

VERBAL INSTRUCTIONS

Close the door five minutes after the scheduled starting time, leaving a DO NOT DISTURB sign and a sign-up sheet on the door.

The sign up sheet states: "We are sorry you were unable to make it to this experimental session on time. So that everybody can finish in a timely manner, we have already begun the experiment. Since you made the effort to participate in this experiment, there will be no penalty for no-showing. Please provide your name and phone number below. Also, please indicate if you would like an opportunity to participate in this experiment at a later date." All such sign up sheets will be identified by date and experimental condition.

Give each participant two copies of the consent form. Provide the following introduction (verbal instructions read by the experimenter are in boldface; taped verbal instructions are in underlined boldface):

"Hi, my name is _____. Thanks a lot for coming today. You were given two copies of a consent form as you came in. If you have not already read it, please go ahead and read it now. Please sign both copies if you have read the form and agree to participate. You should keep one copy for your records and give one copy back to me. Let me know if you have any questions."

Allow a moment for participants to read the form.

"Any questions about the consent form?" After questions: Collect consent forms. "We'll be collecting data later in the semester so its REALLY important that you not talk to other students about this experiment because they may be participating in this study at a later date."

"I'll be playing part of the instructions today on a tape to make sure that each group of students participating in this experiment hears the instructions exactly the same way. I'm going to go ahead and start the tape now."

Start tape: "I appreciate your coming today. In this experiment, we are trying to find out how easy it is for people to fake depression on the Minnesota Multiphasic Personality Inventory, also called the MMPI-2 for short. The MMPI-2 is the most commonly used personality test in the United States and is an update of an older version of the same test. As with any revision of a psychological test, we need to do research to determine how well the test is functioning. A lot of significant research on the original

MMPI was done right here at Texas Tech through experiments with students such as yourselves.

Today I will be asking you to take on the role of a depressed person and to fill out the MMPI-2 as if you are actually depressed. In the real world, most people who take the MMPI-2 answer in an honest manner. However, there are some people who respond in dishonest ways. For example, a criminal might fake a serious mental disorder on the MMPI-2 so that he or she could be placed at a psychiatric hospital instead of a maximum security prison. Therefore, it is important to know how susceptible the MMPI-2 is to people who are trying to fake mental illness. By participating in this experiment, you will be helping psychologists to better determine when people are responding dishonestly. Therefore, it is really important that you put effort into your performance so that the results we find will be meaningful. I realize that some of you may find the idea of lying about the way you are feeling offensive, but it will be extremely helpful to know if you can accomplish what we are asking you to do.

Stop the tape and say "Does anybody have any questions?"

Answer any questions.

Experimenter should then say: "Now I'm going to hand out some materials to you. I'd like you to wait to open the packet until I tell you to do so. After handing out the packets, say, 'I'm going to read off everyone's name and when I get to yours, please tell me the number on the top right hand corner of your packet.'

Record the number next to each name and verify that everyone present is on the list and has an assigned number.

"You can go ahead and open up your packets now. Please pull out the bunch of white papers that are stapled together that look like this (hold up a sample). You can leave the rest of the materials in the manila envelop. I'd like you to look at the first sheet called 'Demographic Information' and for you to fill out the items you see there. Please do not turn to the next page until I tell you to do so."

When this is accomplished, say, "Please turn to the next page and follow along with the tape I am about to play. As you read the story, try to imagine that you are actually experiencing these events as you read them. When you have finished, please do not turn to the next page."

PLAY TAPE OF SCENARIO.

Tape continues: "As you can see, we want you to imagine that you are the student that I just described to you. We want you to take the role of a person who is trying to fake depression on the MMPI-2."

Stop the tape and say "Does anybody have any questions?"

Answer all questions as concisely as possible, referring to the scenario whenever possible.

"As you read on the consent form, cash prizes will be awarded to the top 5 fakers of depression who participate in this experiment. The number one person will get \$75, the number two will get \$30, and so on. When you have completed the experiment, I will ask you if you want to be eligible for these prizes. If you do, I will ask you to leave your phone number so that I (if you are not Gina, say "the principal investigator" rather than "I") will have some way to contact you if you win a prize. Any questions about this?"

Also, please do not talk to your friends or your classmates about this experiment or the cash prizes involved. Only participants in certain conditions are eligible to win prizes. So, ya'll happen to be in one of those conditions, but some of your classmates may not be. Any questions about that?

"Ok, I'm going to start the tape again."

THE FOLLOWING SECTION IS FOR THE VALIDITY INFORMATION GROUP OR THE VALIDITY AND DEPRESSION GROUP ONLY!

Tape continues: "In order for you to better role-play a person who is trying to fake depression on the MMPI-2, we are going to give you some information about the psychological test you are about to take. We want you to learn about some of the ways that the MMPI-2 can tell if someone is not answering honestly. So, turn to the next page and follow along with me as I read the sheet titled 'Information on the MMPI-2 Validity Scales.'"

PLAY TAPE OF VALIDITY INFORMATION

After this, the experimenter should say: "On the next page you will find a brief worksheet to help you learn this material. I'm going to give you about 10 minutes to fill it out. Please feel free to refer to the information sheet we just read. It's REALLY important that you work carefully because you will be quizzed over this information in a few minutes."

After about 10 minutes, or after everyone appears to have finished filling out the worksheet, experimenter should say, "Looks like all of you are about done. Now I'm going to hand out the answers to the worksheet. Please take a few minutes to check your answers."

(PASS OUT VAL SCALE WORKSHEET ANSWERS sheet).

After a few minutes, state, "OK, now I'd like you to pass the worksheet answer key up to the front." Collect the

materials and then say, "Now I'm going to pass out the quiz on the validity scales. Please leave it face down on your desk until I tell you to begin." After passing out the quizzes say, "We are quizzing you about the information you just learned to make sure you understand it. You'll have about five minutes to answer the 10 questions on the quiz. Be SURE to write the i.d. number listed on the outside of your packet in the space in the top right hand corner of the quiz. Please make sure it is legible. When you finish, just bring the quiz up to me. PLEASE DO NOT refer to the validity scale information or the worksheet while answering the quiz. You can start now.

***When each participant turns in the quiz, check to make sure the i.d. number has been written legibly on the quiz.

THE FOLLOWING SECTION IS FOR THE DEPRESSION GROUP OR THE VALIDITY AND DEPRESSION GROUP¹⁵

After the taped presentation of the validity scale information outlined in the preceding verbal instructions for groups receiving validity scale information, the experimenter should say: "There is another kind of information we want you to learn that may help you fake depression more successfully on the MMPI-2. Before I start the tape again, please get out the bunch of stapled white papers we were looking at earlier."

Start tape

"We are going to give you some information about the types of symptoms people with depression typically have. So turn to the next page and follow along as I read the sheet titled "Symptoms of Depression."

After the depression information has been played, the experimenter should say: "On the next page you will find a brief worksheet to help you learn this material. I'm going to give you about 10 minutes to fill it out. Please feel free to refer to the information sheet we just read. It's REALLY important that you work carefully because you will be quizzed over this information in a few minutes.

After about 10 minutes, or after everyone appears to have finished filling out the worksheet, say, "Looks like all of you are about done. Now I'm going to hand out the

¹⁵ The verbal introduction before the tape of the depression information was played was slightly altered for the depression group and for the validity and depression group (if depression information was presented before the validity information).

answers to the worksheet you just filled out so that you can make sure you answered the questions correctly. Please take a few minutes to check your answers."

After a few minutes, state, "OK, now I'd like you to pass the worksheet answer key up to the front." Collect the materials and then say, "Now I'm going to pass out the quiz on depression symptoms. Please leave it face down on your desk until I tell you to begin." After passing out the quizzes say, "We are quizzing you about the information you just learned to make sure it made sense to you. You'll have about five minutes to answer the 10 questions on the quiz. Be sure to write the i.d. number listed on the outside of your packet in the space in the top right hand corner of the quiz. Please make sure it is legible. When you finish, just bring the quiz up to me. PLEASE DO NOT refer to the depression information or the depression worksheet while answering the quiz. You can start now."

***When each participant turns in the quiz, check to make sure the i.d. number has been written legibly on the quiz.

THE FOLLOWING SECTION IS FOR ALL FAKING GROUPS

Experimenter should say: "Before I start the tape again, I'd like you to look in your packets and pull out the MMPI-2 Test Booklet and the MMPI-2 answer sheet. (show participants what they look like) Before looking at the test, I'd like you to listen to these taped instructions."

Tape: "Remember this while responding to MMPI-2 Items. You're really in a jam because you are doing poorly at school and are having financial problems. You are about to lose your scholarship and may have to drop out of school. If you can successfully convince the person interpreting the MMPI-2 that you are severely depressed, you can receive financial assistance and help from your professors. So remember that you are trying to fake having depression on the MMPI-2. Feel free to reread the scenario at any time during the test. Filling out the MMPI-2 can take as long as 2 hours, but some people will finish the test quicker than others. After completing the MMPI-2, you will then fill out a short feedback questionnaire which is located in your packet of materials in a manila envelop. When you finish the questionnaire, please place it and the MMPI-2 back into the larger manila envelop and bring it up to the front. At that point, you will receive your experimental credit slip and will be free to leave. Before you begin answering the items on the MMPI-2, be sure to read the instructions written in the front of the booklet. You can go ahead and start now."

****EXPERIMENTER INSTRUCTIONS FOR ALL FAKING GROUPS AFTER
THE MMPI-2 IS IN PROGRESS****

Twenty minutes after beginning the MMPI-2, say, "OK, just a little reminder - Don't forget that you are trying to appear as if you have depression. Remember that your college career may be on the line here." Forty minutes into the test say, "Remember to think carefully about your answers and remember that you are trying to fake depression."

After each participant turns in the packet, MAKE SURE they have filled out the FEEDBACK QUESTIONNAIRE. Then give him or her an experimental credit AND THE DEBRIEFING FORM and thank him or her for participating.

*****ALSO, tell them to leave their phone number if they wish to be eligible for the cash prizes. Have them write their phone number on the sheet with their names and ID numbers.

APPENDIX D.2: VERBAL INSTRUCTIONS
FOR HONEST RESPONDING GROUP

VERBAL INSTRUCTIONS - HR

Close the door five minutes after the scheduled starting time, leaving a DO NOT DISTURB sign and a sign-up sheet on the door.

The sign up sheet states: "We are sorry you were unable to make it to this experimental session on time. So that everybody can finish in a timely manner, we have already begun the experiment. Since you made the effort to participate in this experiment, there will be no penalty for no-showing. Please provide your name and phone number below. Also, please indicate if you would like an opportunity to participate in this experiment at a later date." All such sign up sheets will be identified by date and experimental condition.

Give each participant two copies of the consent form. Provide the following introduction (verbal instructions read by the experimenter are in boldface):

"Hi, my name is _____. Thanks a lot for coming today. You were given two copies of a consent form as you came in. If you have not already read it, please go ahead and read it now. Please sign both copies if you have read the form and agree to participate. You should keep one copy for your records. Let me know if you have any questions."

Allow a moment for participants to read the form.

"Any questions about the consent form?" After questions: Collect consent forms. "We'll be collecting data later in the semester so its **REALLY** important that you not talk to other students about this experiment because they may be participating in this study at a later date."

"Today you will be taking the Minnesota Multiphasic Personality Inventory, also called the MMPI-2 for short. The MMPI-2 is the most commonly used personality test in the United States and is an update of an older version of the same test. As with any revision of a psychological test, we need to do research to determine how well the test is functioning. A lot of significant research on the original MMPI was done right here at Texas Tech through experiments with students such as yourselves."

Does anybody have any questions?"

Answer any questions.

Experimenter should then say: "Now I'm going to hand out some materials to you. I'd like you to wait to open the packet until I tell you to do so.

"You can go ahead and open up your packets now. Please pull out the white piece of paper titled "Demographic Information" and please fill it out. (hold up a sample). You can leave the rest of the materials in the manila envelop.

After participants have filled out demographic sheet:

I'd like you to look in your packets and pull out the MMPI-2 Test Booklet and the MMPI-2 answer sheet. (show participants what they look like). What I'd like for you to do today is to fill out the MMPI-2 as honestly as possible. It may take as long as 2 hours to fill out, but some people will finish the test quicker than others. After completing the MMPI-2, you will then fill out three short questionnaires which are located in your packet of materials in 2 separate envelopes. It is important that you fill out the questionnaires contained in the envelop labeled "#1" first (hold up example). Then you will need to fill out the single questionnaire located in the other envelop which is labeled "#2". When you finish the questionnaires, please place them and the MMPI-2 back into the larger envelop and bring it up to the front. At that point, you will receive your experimental credit slip and will be free to leave. Before you begin answering the items on the MMPI-2, be sure to read the instructions written in the front of the booklet. You can go ahead and start now.

****EXPERIMENTER INSTRUCTIONS AFTER THE MMPI-2 IS IN PROGRESS****

After each participant turns in the packet, MAKE SURE they have filled out the FEEDBACK QUESTIONNAIRE and the TWO QUIZZES. Then give him or her an experimental credit AND THE DEBRIEFING FORM and thank him or her for participating.