THE EFFECTS OF FAKING INSTRUCTIONS ON MMPI-2 RESPONSE LATENCIES

By

DAVID GENE BRUNETTI

Bachelor of Science
Colorado State University
Ft. Collins, Colorado
1987

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1989

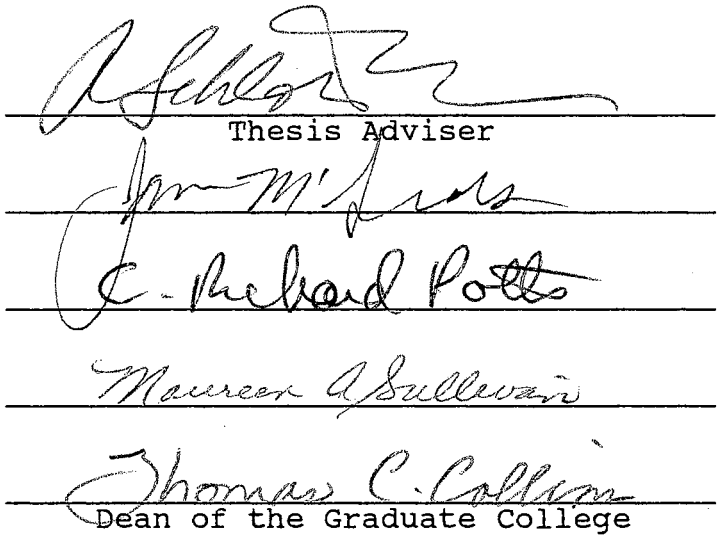
Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 1994

THE EFFECTS OF FAKING INSTRUCTIONS

ON MMPI-2 RESPONSE

LATENCIES

Thesis Approved:



Thesis Adviser

Dean of the Graduate College

ACKNOWLEDGEMENTS

I want to express my gratitude to Robert Schlottmann, Ph.D. for his support, guidance, and persistence in completing this project and the Ph.D. program. I extend my thanks the other members of my committee, Richard Potts, Ph.D., James Seals, Ph.D., and Maureen Sullivan, Ph.D. for their encouragement and feedback. In addition, I want to thank the Department of Psychology as a whole for the opportunity to pursue an advanced degree, and Daniel McNeil, Ph.D. in particular for his friendship and assistance in developing as a psychologist.

Many people contributed friendship, fun, and love to my life (some are no longer a part of it) during this process. I offer my sincere thanks to my ex-wife Laura Brunetti, Marty Boone, Michael Lewin, Cary Savage, Leslie Schwarz, the whole golfing group, the die-hard Willie's crew, and others which I have probably overlooked.

Finally, and most importantly, I thank my parents, Larry and Virginia, my brother Jon, and my sister Tina, for their unwavering love, support, and understanding before, during, and after my schooling. You made this possible and I love you all very much.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

The self-report of individuals is important in many domains of research and applied practice in psychology. Psychological testing, in particular, depends very heavily on verbal report. For this reason, clinicians and researchers interested in personality assessment have been concerned with methods of detecting and measuring the accuracy or validity of an examinee's responses to personality inventory items. Test validity refers to the ability of an instrument to predict behavior and/or its relationship to other measures which are related to the behavior of interest. Validity of an individual's personality inventory profile depends upon the accuracy with which an examinee's pattern of responses describes the person's "true" personality.

Assessment of personality and psychopathology is important in research and clinical work. In research, an accurate estimate allows for the control of extraneous variables, the selection of subjects, and determination of experimental outcomes. In clinical settings, assessment is needed for diagnosis, treatment planning, and monitoring of treatment progress.

Collections of techniques (objective and projective tests, behavioral assessments) are frequently employed. The Minnesota Multiphasic Personality Inventory (MMPI, Hathaway

& McKinley, 1967) is one such objective measure of psychopathology. It is currently the most widely used measure in this area (Tarter, 1988). The MMPI was designed with specific scales to determine whether a deviation from accurate reporting, caused by factors such as faking or carelessness, had occurred.

Hathaway and McKinley (1943) developed several validity scales as they constructed the clinical scales. The Cannot Say (?) scale consisted of the number of items omitted or answered both true and false. It provided a check for carelessness and avoidance of specific item content. The Lie (L) scale was designed to indicate a "fake good" pattern (underreporting psychopathology) as well as "acquiescent" (a tendency to answer "true") response styles. The Frequency (F) scale consisted of items that were rarely endorsed by control subjects. The F scale was intended to detect "fake bad" (overreporting psychopathology), tendencies to respond "true," and random response styles. The Correction (K) scale was added, after cross-validation, to increase the predictive validity of several clinical scales (Meehl & Hathaway, 1946; Dahlstrom, Welsh, & Dahlstrom, 1972). In addition, it has been conceptualized as a measure of defensive responding or subtle faking good (Graham, 1987). Gough (1947; 1950) introduced the F minus K "dissimulation index" to detect both types of faking.

In addition to scales specifically designed to assess the accuracy of an individual's pattern of responses,

researchers have examined the "subtle" and "obvious" items (e.g., Wiener, 1948; Wiener & Harmon, 1946) on the MMPI. The subtle and obvious items exist because of the empirical test construction method used by Hathaway and McKinley (1943) in designing the MMPI. Obvious items are easily related to the psychopathology they are intended to measure, whereas subtle items are difficult to interpret with regard to psychopathology. Meehl (1945) hypothesized that obvious items would be easily manipulated by examinees who were faking, but the subtle items would not. Many empirical studies have supported the idea of obvious item susceptibility to manipulation (e.g., Burkhart, Christian, & Gynther, 1978; Peterson, Clark, & Bennett, 1989; Wales & Seeman, 1968, 1969). However, controversy remains about the relationship between item subtlety and the prediction of psychopathology.

Response latencies are the most recent addition to the faking detection literature. Rogers (e.g., 1974a, 1974b) proposed that, when responding to self-referenced items, an examinee must access a self-schema which is hypothesized to be large and complex. However, according to Hsu, Santelli, and Hsu (1989), subjects instructed to fake may access a schema that is different than honest subjects, one that is smaller and less intricate. Hsu et al. (1989) are the only researchers to have examined latencies and faking on the MMPI. Their data indicated that faking groups may take less time answering items than honestly responding groups, and

that characteristics other than the true/false responses examinees give to the MMPI may be useful indicators of profile validity. Given the widespread use of the MMPI (Tarter, 1988), further research must be conducted to explore the use of response latency in detecting faking.

CHAPTER II

REVIEW OF THE LITERATURE

The use of response latencies in the detection of over- and underreporting of psychopathology on the Minnesota Multiphasic Personality Inventory (MMPI, Hathaway & McKinley, 1967) will be examined in this literature review. Several theories have been proposed regarding response latencies when responding honestly to adjective checklists and personality inventory items (e.g., Rogers, 1971; Nowakowska, 1970) and will be examined. While the theories are generally consistent, many of the empirical investigations into response latencies have been contradictory to one another and to latency theories. In addition, there is a paucity of theories and empirical research linking response latencies with the detection of faking, especially using the MMPI.

One of the main questions in exploring response latencies and faking on the MMPI involves how the item latencies will be grouped for analysis. Grouping latencies based on the full MMPI and/or on the individual validity and clinical scales are two obvious options. Subtle-obvious scales (e.g., Wiener & Harmon, 1946) will be discussed in terms of their traditional use in predicting profile validity and will be suggested as a possible strategy for grouping latencies. Patterns of "item endorsement" will be examined. Item endorsement refers to the patterns of

"accepted" and "rejected" items. Accepted items are those which increase the score for a particular scale. Rejected items do not increase scaled scores. Popham and Holden (1990) suggested that a relationship exists between MMPI content scale scores and response latencies. Finally, methodological issues concerning response latencies on the MMPI will be discussed.

Response Latencies and Faking

Some authors (e.g., Groth-Marnat, 1990) have suggested that response latencies are important in personality assessment, both as diagnostic indicators and validity measures. However, the use of response latencies with tests, ranging from the MMPI to the Rorschach inkblot test, has received sporadic research. Beyond clinical observations, very little research has been performed to evaluate the use of response latencies as a validity measure on psychological tests. Several theories have been advanced to explain the hypothesized relationship between latencies and faking. These theories also remain relatively untested, particularly with the MMPI.

Rogers (1974a, 1974b) identified several stages in the process of responding to personality inventory items using Sternberg's (1969) method of examining information processing sequences. Item characteristics such as length, ambiguity, and controversiality were found to increase response time additively, but not multiplicatively.

According to Sternberg, factors which affected reaction time, but did not interact, were indicators of independent sub-processes. Those attributes were hypothesized to reflect the stages of stimulus encoding, stimulus comprehension, and decision-making, respectively (Rogers, 1974a). Rogers (1974b) suggested that the decision-making stage, influenced by item controversiality, was independent of the response format (true/false versus Likert-type format). It was shown that item controversiality and latency were positively related, as were the number of response choices (two or five) and latency. Item controversiality and response format, however, did not show an interactive effect on latency. Rogers concluded that the increase in latency with highly controversial items was not due to difficulty translating a probabilistic decision ("I'm sad some of the time") into a binary response (true or false), but was related to the difficulty of the decision itself. By isolating the decision stage, researchers were able to examine changes in response latency associated with different factors that effect a respondent's decision (Rogers, Kuiper, & Kirker, 1977).

Latencies could theoretically be increased or decreased by changes in any of the item characteristics which affect the above mentioned stages. For example, if highly controversial items were used, the decision stage would be made more difficult, and longer latencies would be expected. Most interest has focused on the decision stage where

instruction sets can be imposed to alter the nature of the decision. Rogers et al. (1977) asked subjects to respond to a list of adjectives in one of four ways. The instructions created a structural decision ("Are the letters big?"), a phonemic decision ("Does the word rhyme with XXXX?"), a semantic decision ("Does the word mean the same as XXXX?"), and a self-referent decision ("Does the word describe you?"). Each type of decision, in the order listed above, theoretically required "deeper" processing and longer response latencies. The effects of depth of processing have been demonstrated in human memory research (e.g., Craik & Lockhart, 1972; Craik & Tulving, 1975; Jacoby & Craik, 1979). Rogers et al. (1977) indicated that response latency and the number of words recalled systematically varied with the instructions an examinee received. A self-referenced decision, theoretically, produces more elaborate processing because it necessitates reference to a "self-schema." Self-schema was defined by Best (1986) as "...a complex cognitive structure into which information about the self is accepted and organized" (p. 85). Specifically, latency and words recalled increased as the "depth of processing" involved in the decision increased, with the self-referenced decision eliciting longer latencies and producing greater recall than the other tasks. Deeper or more elaborate levels of processing have produced improved recognition and recall. While aiding memory, more elaborate processing presumably results in increased cognitive effort and longer response

latency.

Cantor and Mischel (1977) found that subjects rated highly on a personality schema characteristic, such as introversion, recognized more previously presented schema-related than schema-unrelated adjectives. Subjects also demonstrated a "schema bias." They reported recognizing schema-relevant items, which had not been previously presented, more often than irrelevant items which had not been previously presented. As noted above, Rogers et al. (1977) demonstrated the longest latencies, in addition to the greatest incidental recall, with adjectives processed using a self-referenced decision. The authors concluded that different schemas were being accessed by the different instructions, with the self-schema being most useful to memory but requiring the most processing time. Possibly, while facilitating more elaborate encoding, the self-schema slows processing.

Responding to MMPI items with the intent to deceive may involve accessing a less elaborate information network or schema. In faking mental illness or health, the respondent compares the item with a schema for psychopathology that is hypothesized to be much smaller and/or less complex than the self-schema (Hsu, Santelli, & Hsu, 1989). With a less complex network being accessed, shorter response latencies are expected for subjects instructed to faking than instructed to respond honestly (self-referent) responding. Hsu et al. (1989) presented means and standard deviations

for latencies of honest, fake bad and fake good groups administered the MMPI. The authors were primarily interested in the predictive validity of subtle/obvious items versus response latencies and, therefore, provided no formal statistical tests among means. Inspection of the means, however, suggested that subjects faking both good and bad had shorter latencies than subjects responding honestly. Also, subtle items appeared to require more response time than obvious items.

Additional theories have been proposed to describe the relationship between response latency and personality inventory items. These theories focused primarily on honest responding, but may be expanded in exploring the relationship between response latencies and faking. Nowakowska (1970) hypothesized that there were three possible decision "paths" in responding to a test item. The first was based on the content of the item, what the item was actually stating. Social desirability was said to comprise the second path, where answers were appraised for negative evaluation from others. Acquiescence or preference for specific types of answers, such as true or false, made up the third path. Feedback loops were hypothesized to check the truth and utility of output of three possible responses before an actual response was given. Under faking instructions the first and third paths would likely be devalued if not eliminated (Hsu et al., 1989), since faking is heavily influenced by social desirability. The result of

only one path being utilized would suggest a decrease in response latency for faking over honest answering, due to a decrease in the amount of processing necessary to respond.

Kuncel (1973) theorized that latency was a function of the "distance" between the examinee's "decision threshold" on an issue and the position of the item. An example from her article involves the legal voting age (pp. 545-546). If a subject is asked whether people varying in age between 10 and 30 should be able to vote, a fairly quick response is expected with the 10 - 15 year range and the 25 - 30 year range. However, the decisions around 18 to 21 years old are likely to involve more processing time. Kuncel stated that most examinees' attitudes or positions on the legal voting age lie near the 18 - 21 year range. Decisions on those items are difficult, according to Kuncel, because being nearest the subjects' own threshold on the issue, many factors must be weighed in order to respond. Items more distinct (farther away) require fewer factors to be examined. In other words, items which were very similar to the respondent's own threshold were expected to generate longer latencies than dissimilar items.

Kuncel's (1973) model was intended to explain responses from subjects instructed to answer honestly. Three assumptions must be made if her model is extended to subjects instructed to fake on the MMPI (Hsu et al., 1989). First, MMPI items are assumed to be positioned near the middle of a continuum of mental health. Applying this

assumption to the MMPI is difficult to defend because many of the items are obviously related to psychopathology and would be nearer the mental illness extreme. Second, the position of a "normal" subject responding honestly would be somewhere between the item, assumed to be near the middle of the continuum, and a fake good position. Third, faking instructions shift in a person's usual position on a continuum of mental health to the extreme. The examinee's position moves to the mental illness extreme when instructed to fake bad and to the mental health extreme when instructed to fake good. The result would be that faking instructions move an examinee's position farther from the item than honest instructions. Given those assumptions, a normal subject responding honestly would occupy a position closer to the item than a subject instructed to fake, who would be at the extreme. The result would be longer latencies for honest subjects compared to faking subjects. Longer latencies for honestly responding subjects is consistent with the expectations arising from Rogers (e.g., 1977) and other researchers.

Several studies have produced results which are inconsistent with the theoretical discussions and empirical findings of Rogers (e.g. 1977), Nowakowska (1970), and Kuncel (1973). Markus (1977) examined differences in response latencies on self-referenced adjective checklists for female college students. The subjects were separated into one of three groups: those who indicated a self-schema

for independence ("independents"), a self-schema for dependence ("dependents"), or no self-schema for independence/dependence. She found significant between- and within-groups differences. Dependent subjects accepted adjectives related to dependence faster than independent subjects, while independent subjects accepted adjectives related to independence faster than dependent subjects. Within-groups, dependent subjects were faster responding to dependent adjectives than independent adjectives, whereas the opposite was true of independent subjects. The result was a positive linear relationship between schema-related items and response latency; items most closely associated with an individual's self-schema resulted in shorter latencies than distantly associated items. These results suggested that a self-schema facilitated processing of the schema-relevant adjectives, contrary to Rogers and others.

Other researchers have demonstrated findings contradictory to Rogers. Kuiper (1981) had undergraduate students rate themselves, using a 9-point scale, on 24 adjectives from the Personality Research Form (PRF; Jackson, 1967). Items that were rated most and least self-descriptive produced the shortest latencies while the moderately self-descriptive items had the longest latencies, an inverted-U curve. Additionally, Kuiper (1981) presented data from Kuiper and Rogers (1979) in re-analyzed form and demonstrated the inverted-U latency curve again with self-referrent judgements.

Signal detection theory has also been proposed as an explanation for responses to personality test items. Hanley (1962) investigated response time on "controversial" MMPI items, defined as those items which have an approximately equal probability of being answered true or false by normal subjects. Hanley demonstrated longer reaction times for controversial MMPI items than noncontroversial items.

Rogers (1971) proposed a signal detection theory of responses to personality test items which helped clarify controversiality and MMPI items. Rogers (1971) stated that an individual's response to an item was determined by the interaction of the social desirability of the item, the "salience" of social desirability to the individual, and the individual's social desirability "threshold." The threshold was defined as the critical level of social desirability needed to make the transition from a "true" response to a "false" response. Items close to the threshold have approximately an equal chance of being answered true or false and responses to those items are therefore longer than items which fall farther from the threshold. Rogers (1971) stated that if salience was high for a subject then social desirability would be a major determinant in that subject's responses, but if salience was low then other factors, such as item content, would determine the responses. The introduction of faking instructions would certainly affect an individual's threshold for desirability and the salience of social desirability. For example, if a college student

were instucted to over-report symptoms of psychopathology, both the salience of pathology-related items and the subject's probability of answering "true" to pathology items would increase.

MMPI Item Subsets and Response Latency

A relationship between faking and response latencies on personality tests could only be inferred from the theories and empirical studies reviewed above. None of the researchers included faking instructions in their designs, so conclusions about faking are tenative. In addition, all the researchers used adjective checklists rather than MMPI items. Adjective checklists are significantly different from the MMPI items in form (single word vs. sentence) and response format (Likert-type scale vs. true-false response).

Another major difficulty to be overcome in studying latencies on the MMPI involves how the item latencies are grouped or which item subsets are compared. For example, latencies for different instruction groups could be compared across the total MMPI as well as for the validity and clinical scales. These comparisons have not been made in any of the MMPI faking studies. In addition, item subtlety and item endorsement are potential strategies for grouping MMPI items in the examination of faking.

Item Subtlety

The MMPI was designed with specific indices, such as

the L, F, and K scales, to determine whether a deviation from honest responding had occurred. In addition to scales specifically created to assess the validity of an individual's pattern of responses, researchers have suggested using the MMPI "subtle" and "obvious" items in detecting faking (e.g., Greene, 1980, 1988).

Hathaway and McKinley (1940, 1943; Butcher, 1990; Graham, 1987) constructed many of the MMPI subscales using the empirical, or criterion keying, method. That approach was chosen due to problems with earlier inventories developed according to the logical keying or face valid procedure. One of the problems with face valid tests involved the ability of an individual subject to fake mental illness or health through patterns of responses. The ease of faking was due to the obvious relationship between inventory items and the psychopathology they were intended to measure. The empirical keying methodology was an attempt to minimize that problem. A large number of items representing a wide variety of content areas were generated and responses to them were collected from two or more criterion groups. Item responses of the criterion groups were compared, and those items that differentiated the groups were retained on the inventory (Graham & Lilly, 1984). For example, Hathaway and McKinley (1942) constructed the Depression (D) scale on the MMPI by comparing a group of clinically depressed patients with a group of control subjects. Items that predicted membership

in the depressed group were retained for the D scale. In this way, statements that directly addressed psychopathology ("obvious") were included. Statements which discriminated the groups, but were not obviously related to psychopathology ("subtle"), were also included.

Wiener and Harmon (1946; Wiener, 1948) devised subtle and obvious subscales for the MMPI to investigate the usefulness of those items. Subtlety subscales were derived for the Depression (D), Hysteria (Hy), Psychopathic Deviate (Pd), Paranoia (Pa), and Hypomania (Ma) scales based on their personal ratings of item subtlety (Wiener & Harmon, 1946). These subscales were not developed for the Hypochondriasis (Hs), Psychasthenia (Pt), or Schizophrenia (Sc) scales because those scales contained too few subtle items in Wiener and Harmon's ratings; they were almost exclusively obvious scales. Also, subscales were not developed for the Masculinity-Femininity (Mf) scale or the Social Introversion (Si) scale. In 1948, the Mf scale had not been shown to reliably predict psychopathology and the Si scale had not yet been added to the MMPI, thus no subscales were included.

Christian, Burkhart, and Gynther (1978) reported subtlety ratings for all the items on the MMPI. They instructed undergraduate students to rate each item on its relationship to psychopathology using a 5-point rating system. Those researchers found sufficient numbers of subtle items for the D, Hy, Pd, Pa, and Ma scales. The Mf

and Si scales were also reported to contain a sufficient number of subtle items and, furthermore, ranked as the two most subtle scales. Christian et al. (1978) reported a significant correlation (r = -.78) between ratings of item obviousness and social desirability. Additional divisions of subtle and obvious items on the MMPI have been produced (e.g., Duff, 1965; Meehl & Hathaway, 1946; Seeman, 1952). Dubinsky, Gamble, and Rogers (1985) concluded that with each attempt to devise subtle/obvious item sets slightly different results were achieved, yet "the practical results are remarkably similar" (p. 67). Regardless of the subtle/obvious item divisions used, subjects produced relatively comparable subtle and obvious score patterns.

Early in the history of the MMPI, several researchers (e.g., Meehl, 1945; Wiener, 1948) suggested using item subtlety in the detection of deviant response sets and in the prediction of personality characteristics when faking occurred. The obvious items were thought to be easily manipulable, due to their face validity, and inaccurate under faking conditions. Subtle items were believed to remain relatively unchanged whether an examinee responded honestly or faked. Therefore, in the event of attempted deception, it was hypothesized that subtle items would give a more accurate description of the examinee's psychopathology than obvious items.

To address the question of item sensitivity to faking, researchers have instructed groups of subjects to take the

MMPI honestly, underreport and/or overreport psychopathology. In general, "fake bad" (overreporting) subjects are expected to produce higher obvious scale scores than honestly responding subjects, while subtle scale scores are not expected to differ between the groups. The converse is expected for "fake good" (underreporting) subjects, whose obvious scale scores should be lower than honest subjects. Again subtle scores are not expected to differ between fake good and honest subjects.

Numerous empirical investigations have demonstrated the ability of subjects to manipulate obvious items on the MMPI while subtle items have shown relative resistance to response sets. Wales and Seeman (1968, 1969) reported that college students and psychiatric inpatient groups instructed to fake good on the MMPI endorsed significantly fewer obvious items than did honestly responding groups. Burkhart, Christian, and Gynther (1978) showed that undergraduate examinees successfully manipulated the obvious MMPI items in the desired direction when instructed to under- and overreport psychopathology. Grow, McVaugh, and Eno (1980) reported the expected changes in obvious items with fake bad and fake good conditions versus an honest condition for student, outpatient psychiatric, and inpatient psychiatric groups. Those results were important because they indicated comparable approaches to faking in psychiatric and nonpsychiatric groups.

Peterson, Clark, & Bennett (1989) demonstrated that

counseling psychology graduate students responding to fake bad instructions were able to manipulate the obvious and full scale scores but not the subtle scales. Under fake good instructions, the same subjects were unable to significantly distort the full, obvious or subtle scale scores. One explanation offered was that normal subjects demonstrate a "floor effect." Normal subjects who respond honestly endorse very few obvious items in the pathological direction. When normal subjects are instructed to fake good, they also endorse few obvious items in the pathological direction. Therefore, differences between the honest and fake good conditions with normal subjectes may be small. The floor effect demonstrated by Peterson et al. (1989) contradicted Grow et al. (1980), who found normal groups able to successfully fake good. Many more studies (Anthony, 1971; Harvey & Sipprelle, 1976; Vesprani & Seeman, 1974; Wales & Seeman, 1972) have shown that obvious items are more easily manipulated than subtle items, a result that has been generally accepted in the literature (Dubinsky et al., 1985; Worthington & Schlottmann, 1986).

Many researchers (Anthony, 1971; Burkhart et al., 1978; Harvey & Sipprelle, 1976; Vesprani & Seeman, 1974; Wales & Seeman 1968, 1969) have demonstrated a "paradoxical" effect under faking instructions, where subtle scores change in a direction opposite to that expected. For example, under fake good conditions the obvious scale scores decrease, as hypothesized, but the subtle scale scores increase.

Conversely, under fake bad conditions the obvious scores increase but the subtle scores decrease. The movement in the subtle scores may serve a corrective function on full scale scores because full scale scores are a combination of subtle and obvious scales. An examinee faking good may successfully decrease the obvious item score while paradoxically increasing the subtle item score. The overall effect may be a full scale score higher than if the subtle items had been totally unaffected by faking. Changes in the subtle item scores, however, are small relative to the changes in obvious item scores.

Only one study has used standardized faking instructions and item response latencies in the detection of dissimulation on the MMPI. Hsu et al. (1989) compared the predictive validity of response latency scores and "response deviance" (subtle-obvious) scores. These researchers used the subtle and obvious items identified by Christian et al. (1978). Subjects were instructed to answer items truthfully, fake good, or fake bad. Latencies were transformed using a log function due to homogeneity of variance problems. The results were not directly interpretable, particularly for clinical use, because of the log-transformed data. Using a canonical correlation procedure, Hsu et. al. (1989) showed that the response latency method had significant predictive validity independent of the scores on subtle-obvious items, or what the authors called "incremental validity." Latencies aided

in the detection of both fake good and bad response patterns. Response latencies enhanced detection of fake good styles more so than fake bad styles. The response deviance scores facilitated the detection of fake bad better than fake good. Additionally, as noted above, their results suggested that faking examinees respond more quickly than honestly responding subjects, although no statistical tests related to this question were presented.

Item Endorsement

Item endorsement refers to the patterns of items "accepted" and "rejected" by an examinee. Popham and Holden (1990) examined the relationship between response latencies and specific MMPI scale scores using patterns of item endorsement. Utilizing factor-analytically derived MMPI content scales (Costa, Zonderman, Williams, & McCrae, 1985), Popham and Holden demonstrated negative correlations between scale T-scores and mean latencies for accepted items and positive correlations between scores and latencies for rejected items. In other words, if an individual scored high on a scale, the items indicated to be representative of the person (accepted items) would have shorter latencies than items unlike the person (rejected items). For low scoring individuals, the pattern is reversed where items similar to the person had longer latencies than dissimilar items. Thus, an examinees' scores on a scale, such as Extraversion (Costa et al., 1985), could be predicted by

examining latencies for endorsed or rejected items (higher scores related to more rapid endorsement). As with earlier experiments (e.g., Dunn, Lushene, & O'Neil, 1972), no relationship was discovered between latencies and the traditional clinical scales. One possible explanation may be that five of the eight clinical scales contain many subtle items which may be less prevalent in the factor-analytic content scales, although no information about the subtleness of the factor-analytic scales was provided. Popham and Holden reported that no previous MMPI research had established a relationship between response latencies and any specific scale scores.

Dunn et al. (1972), using a step-wise regression procedure, reported that the mean latency for admissions of psychopathology on Grayson's (1951) critical items was significantly longer than the mean for rejections of those items. The differences on Grayson's items could be interpreted as support for the hypothesis by Kuiper (1981) and Markus (1977) that items which are related to oneself require less time than unrelated items. Undergraduate subjects were utilized by Dunn et al. (1972) and those subjects are assumed to be "normal." Acceptance of critical items, which are generally psychopathology-related, would be assumed to be schema-unrelated for undergraduate students, thus producing longer latencies.

A different interpretation of the Dunn et al. (1972) data is suggested by the results of Temple and Geisinger

(1990). These researchers constructed a 48-item inventory using the MMPI, Personality Research Form (Jackson, 1967), and specially generated items that were "emotionally evocative" (p. 290) or neutral. The evocative items were found to take significantly more time to respond to than the neutral items. If the critical items were assumed to be "evocative" to undergradute subjects then those items would require longer response times. One generalization to the full MMPI may be that an individual who scores high on specific scales would produce longer latencies than on the lowest scales. The researchers would have to assume, however, that the items on high scales were more arousing than items on low scales. This assumption may be particularly true of obvious items, where endorsement is the admission of socially undesirable psychopathology. The prediction of longer latencies associated with high scale scores, however, is contradictory to the findings of Popham and Holden (1990). Those researchers found shorter latencies for accepted items associated with the highest scale scores.

Methodological Issues

The theoretical literature supports the empirical exploration of the relationship between response latencies and profile validity. Although some of the empirical studies suggested a relationship between latencies and faking, all had at least one methodological deficiency or

could not be applied to faking detection without some assumptions. Hsu et al. (1989) used a combination of item subtlety and latencies on the MMPI, but transformed their data and used a subtle item set (Christian et al., 1978) not typically used by clinicians. As noted above, many of the empirical investigations did not use faking instruction sets and their results had to be extrapolated with regard to faking. In addition, a number of researchers used adjective checklists, which obviously differ from the sentence format used on the MMPI. The revised and renormed MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) has become available and the effects of faking on that instrument must be assessed.

Characteristics of individual MMPI items must considered and controlled if response latencies are to be used as a dependent variable. Dunn et al. (1972) measured MMPI response latencies and found that 47% to 58% of the variance in latencies was due to item length. A much smaller amount of variance was related to item ambiguity and social desirability (3% - 6%). Dunn et al. (1972) suggested that any investigation into response latencies on the MMPI must take item length into account. Popham and Holden (1990) attributed part of their positive results to a *Z*-score standardization procedure. The latency for every item was adjusted for the mean and standard deviation of the individual's items latencies as well as the mean and standard deviation for latencies for each item across

subjects. The Z-score standardization emphasized variance related to decision-making processes while adjusting for subject characteristics, such as reading ability, and item characteristics, such as item lenght and ambiguity.

The review of the literature indicated that much more research is needed to establish which set(s) of items should be combined with response latencies to assist in the detection of faked MMPI profiles. The theories which relate faking and latency must also be examined and revised by empirical results.

CHAPTER III

STATEMENT OF THE PROBLEM

Theoretically, response latencies on the MMPI may be affected by a subject's attempt to fake. Rogers, Kuiper, and Kirker (1977) and others (e.g., Kuncel, 1973; Nowakowska, 1970) proposed response latency theories of responding to personality items. However, their theories did not specifically address faking and none of their empirical studies employed faking conditions. The absence of faking instructions made conclusions about the effects of faking on latencies very tentative. Hsu, Santelli, and Hsu (1989) expanded their theories and predicted that honestly responding subjects should produce longer latencies than subjects who were instructed to fake. Hsu et al. (1989) based their predictions mainly on the theories of Rogers et al. (1977), Kuncel (1973), and Nowakowska (1970). Hsu et al. (1989) presented means and standard deviations for honest and faking subjects on the MMPI, in addition to canonical correlations. No formal statistical tests of the means were used, but inspection of the means suggested faking subjects may have responded faster than subjects instructed to answer honestly.

Several researchers have proposed theories and presented empirical results which contradict the idea that faking speeds processing of personality test items. Markus (1977) found that a self-schema facilitated processing of

schema-relevant adjectives, which suggested that subjects instructed to respond honestly would be faster than those instructed to fake. In addition, Kuiper (1981) found that adjectives which were most and least self-descriptive produced the shortest response latencies. The research of Kuiper (1981) and Markus (1977) supported the idea that a self-schema facilitated processing of self-related material. Finally, the results of Hanley (1962) and Rogers' (1971) signal detection theory suggested that faking may alter an examinee's perception and willingness to respond to socially desirable and undesirable items. Therefore, while a theoretical basis may exist for the use of latencies in faking detection, the sparse empirical literature has provided only minimal and somewhat contradictory support for the use of response latencies for the detection faking.

The main question in studying examinees' response latencies on the MMPI is how the latencies are to be grouped for analysis. One method could be to test for differences among honest and faking subjects for total MMPI response latency, all item latencies added together for each group. Another method may be to look at differences among instruction groups on the validity and clinical scales. A third method for grouping latencies could include using subtle and obvious items. Subjects respond differently to the MMPI subtle and obvious items under faking and honest instruction sets. Meehl (1945) and Wiener's (1948) hypotheses that subtle items were more resistant to faking

than obvious items were supported by many experimenters (e.g., Burkhart, Christian, & Gynther, 1978). Subtle items were hard for subjects to relate to psychopathology and, thus, more difficult to manipulate. Obvious scores were shown to increase under fake bad instructions and decrease with fake good. The difference in scale scores for subtle and obvious items under faking instructions may also be reflected in response latency. Unfortunately, rigorous examination of latency differences among faking and honest examinees has not occurred. As mentioned above, Hsu et al. (1989) presented data which may suggest differences. The means for subtle and obvious items presented by those researchers appeared to be different; no statistical tests were performed. Examinees seemed to respond slower to subtle items than obvious items.

A final method of grouping MMPI response latencies involves "item endorsement" (i.e., whether one responds in the same manner in which an item is keyed). Popham and Holden (1990) instructed subjects to respond honestly on the MMPI. They found that individuals scoring high on a scale produced shorter latencies than low scoring individuals. These researchers demonstrated a relationship between response latencies and scores on MMPI content scales. No previous researchers had reported such a relationship. Therefore, response latencies in conjunction with item acceptance/rejection may be useful in detecting faking.

Several methodological considerations were exposed by

the existing research on item subtlety and response latency. Many of the empirical studies used adjective checklists, rather than the MMPI. The MMPI is currently the most widely used objective personality measure (Tarter, 1988) and improved faking detection would have wide application. The MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) was recently introduced and is probably a more appropriate instrument, because it appears to be very similar to the original MMPI and is expected to replace it. No studies of latency and faking on the MMPI-2 have yet appeared in the literature.

The Wiener-Harmon (1946) subtle/obvious items are the only measures of item subtlety available to most clinicians. For that reason, those items should be used in examining the effects of faking on item subtlety and latency. In addition, this choice is appropriate given the relatively small difference in item subtlety subscales produced by many different researchers (Dubinsky, Gamble, & Rogers, 1985).

Finally, item characteristics must be controlled. These include item length, ambiguity, and controversiality. Dunn, Lushene, and O'Neil (1972) showed that 47% to 58% of the variance in latencies on the MMPI was due to item length; 3% to 8% of variance was related to item ambiguity and social desirability. Popham and Holden (1990) used a norm group and $\underline{Z}$-score standardization procedure to control for item characteristics. Subject characteristics, such as an individual's reading ability, were also controlled by

Popham and Holden. In the current study, variance contributed by subject characteristics was assumed to be small since a homogeneous subject group (first-year undergraduates) was used and random assignment to groups was expected to evenly distribute those characteristics.

The hypotheses of this research project were made with knowledge of the lack of consistent empirical results and, consequently, were all nondirectional. First, differences in response latencies were predicted among groups instructed to respond honestly, to overreport, and to underreport psychopathology. Second, the obvious and subtle scales were expected to produce differences in response latencies. Third, differences in latency were anticipated for the accepted and rejected items across the three instruction groups. Finally, the two- and three-way interactions of instruction group, item subtlety, and item endorsement were examined for differences in response latencies.

CHAPTER IV

METHOD

Subjects

One hundred thirty-seven undergraduate introductory psychology students were used as subjects. Twenty-seven (9 females, 18 males) were excluded from data analysis for one or more of the following reasons: invalidity under honest instructions (L > 59, F > 79, or K > 64; N = 16), failure to follow instructions (N = 5), language difficulties (N = 4), or equipment failure (N = 1). Sixty subjects were unsystematically assigned to one of three gender-balanced instruction groups: honest, fake bad, or fake good. The remaining 50 (25 female, 25 male) subjects formed a norm group for standardization of the instructed groups' response latencies. One extra credit point per hour of participation was awarded.

Instruments and Apparatus

The Minnesota Multiphasic Personality Inventory - 2 (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) was the only instrument used. Scores for the three validity scales (L, F, K) and the ten clinical scales (Hs, D, Hy, Pd, Mf, Pa, Pt, Sc, Ma, Si) were computed. The Wiener and Harmon (1946) subtle and obvious subscales of the D, Hy, Pd, Pa, and Ma scales were also calculated. Only the first 370 items of the MMPI-2 were needed to score the above

scales.

The MMPI-2 items are protected by copyright laws and were presented with permission of the University of Minnesota (see Appendix A). Items were displayed on a 12-inch amber monochrome monitor controlled by an IBM-compatible computer and BASIC program. The computer also recorded true/false responses and gathered response latencies (interval between item presentation and subject response). Subjects responded by pressing designated keys on the computer keyboard. A second BASIC program computed Z-scores for each individual's 370 latencies by subtracting the corresponding mean latency derived from the norm group and dividing the result by the corresponding standard deviation. For example, the mean latency to item 1 for the norm group was subtracted from each individual's latency to item 1; the result was divided by the norm group's standard deviation for item 1, producing a Z-score for that item. After each individual's latencies were standardized by the norm group, a third BASIC program calculated validity, clinical, and subtle-obvious T-scores and mean response latencies for each scale. In addition, mean response latencies were calculated for accepted items, rejected items, and total (accepted plus rejected) items on each of the validity, clinical, and subtle-obvious scales listed above.

Procedure

All subjects participateed individually in the study. Responses were identified only by identification number and, thus, remained confidential. Subjects read and signed an informed consent statement before any data were collected (see Appendix B). Items were presented individually on the monitor and subjects were required to press the space bar after each response to display the next item. Each subject was instructed on how to respond to items (see Appendix C). A "1" indicated "true" and a "2" indicated "false." Subjects were not able to respond "cannot say" or review their answers, as allowed in the standard version of the MMPI-2. Reviewing responses would invalidate the latency associated with the true/false answer and a "cannot say" option would further complicate the item endorsement question. Honaker (1988), in a review of computerized and conventional MMPI administration, discussed programs which omitted either the review option, the "cannot say" option, or both. Honaker stated that MMPI profiles from programs which omit both options were most likely to be "nonequivalent" with conventionally-administered MMPI profiles. The potential nonequivalence was not problematic, however, because no comparisions were being made to traditionally-administered MMPI profiles.

The instruction set a subject received depended on the group to which he/she was assigned. All instruction sets were presented on the computer and were slightly modified

versions of those used by Hsu, Santelli, and Hsu (1989; see Appendix). Subjects in the norm and honest groups were given the following instructions:

> I am interested in the psychological characteristics of the student body at this college and I want you to take a personality test for me. Because I am interested in group characteristics I do not want you to [enter] your name [with your responses.] Do, however, accurately record your [gender]. In short, I want you to take this test in an honest but anonymous fashion. (p. 282)

Fake bad subjects received these instructions:

> I want you to imagine a situation in which you are being cast in an adversary relationship against a psychologist. More specifically, imagine a situation in which it would be to your advantage to appear as if you [were mentally disturbed]. Examples of such a situation could be: applying for rehabilitation services, trying to qualify for disability benefits, or trying to beat a legal charge on grounds of insanity. In short, I want you to take this test and deliberately fake bad, so that your deception could not be detected by a professional psychologist. (p. 282)

For fake good subjects the instructions were:

> I want you to imagine a situation in which you are being cast in an adversary relationship against a psychologist. More specifically, imagine a situation in which it would be to your advantage to appear as if

> you were completely normal and sane. Examples of such a situation could be: trying to secure an early release from prison, trying to secure a release from a mental hospital, or applying for a good job. In short, I want you to take this test and deliberately fake good, so that your deception could not be detected by a professional psychologist. (p. 282)

After completion of the MMPI-2, subjects were debriefed with regard to the intention and purpose of the study. Only information about the group in which a subject participated was revealed; a subject who received honest instructions was not told about the faking groups. This procedure was instituted to ensure that future subjects did not become aware of other instruction sets and bias their performance. No feedback on MMPI-2 scores was given. All subjects were given the phone number and address of psychological clinics located on the university campus both and were encouraged to contact either service if items on the MMPI-2 triggered distress.

Data analysis

An initial analyses of variance (ANOVA) was conducted to assess whether the instructions were effective. Standard K-corrected T-scores for the validity and clinical scales were analyzed across the three instruction groups. The fake bad group was expected to produce higher scaled scores than the honest group and the fake good was expected to show

lower scaled scores than the honest group.

One of the BASIC programs computed means and standard deviations for latencies for each of the 370 MMPI-2 items across all 50 norm group subjects. Individual latencies for all subjects which were shorter than 1s or longer than 20s were set at 1s and 20s, respectively. These cutoffs were established using a Statistical Analysis System (SAS) UNIVARIATE procedure and were used to control for the effects of statistical outliers. Popham and Holden (1990) reported that cutoffs of 0.5s and 40s resulted in less than one percent of all latencies being adjusted. The 1s and 20s cutoffs used in the current study caused less than two percent of the raw item latencies to be adjusted.

The means and standard deviations for latencies of the norm group were also used to compute Z-scores for each honest, fake good and fake bad subject's individual latencies, producing 370 Z-scores for each subject. This procedure controlled for item characteristics, such as length, and average university reading ability in the three instruction groups. After standardization, mean Z-scores for latencies were calculated for the 13 standard MMPI-2 scales and the 10 Wiener/Harmon subtle and obvious scales. Mean Z-score latencies were also computed for accepted and rejected items across the validity and clinical scales as well as the subtle/obvious subscales.

Two additional ANOVAs, both with response latency Z-scores as the dependent measure, were conducted. All

subsequent analyses were completed using simple effects ANOVA's and Tukey's post-hoc tests. The first response latency analysis was a 3 X 2 X 13 X 2 ANOVA with two between-subject factors (instruction groups, gender) and two within-subject factors (scales, endorsement). The second response latency analysis was a 3 X 2 X 5 X 2 X 2 ANOVA with two between-subject factors (instruction groups, gender) and three within-subject factors (scales, endorsement, item subtlety).

CHAPTER V

RESULTS

Results of the initial analysis of variance (ANOVA) for instruction groups X scales, using *K*-corrected *T*-scores as the dependent variable, suggested that the fake bad, fake good, and honest instructions differentially affected scaled scores, $F(24,684) = 57.18$, $p < .0001$. Simple effects ANOVA's among groups at various levels of scales and Tukey's multiple comparisons tests revealed that subjects instructed to fake bad scored higher than those instructed to fake good or respond honestly on all scales but *L* and *K*, where they scored significantly lower. Significant differences were noted between subjects instructed to fake good and respond honestly on the *L*, *F*, *K*, and *Si* scales. Table I provides detailed information on the differences among groups.

TABLE I

MEAN K-CORRECTED T-SCORES AT LEVELS OF INSTRUCTION GROUP AND SCALE

	Instruction Group		
	Fake Bad	Fake Good	Honest
L	49[a]	70[b]	48[a]
F	120[a]	45[b]	52[c]
K	37[a]	61[b]	47[c]
Hs	81[a]	52[b]	53[b]
D	72[a]	45[b]	49[b]
Hy	70[a]	49[b]	49[b]
Pd	80[a]	46[b]	51[b]
Mf	63[a]	52[b]	51[b]
Pa	107[a]	50[b]	51[b]
Pt	86[a]	50[b]	52[b]
Sc	109[a]	51[b]	52[b]
Ma	82[a]	53[b]	56[b]
Si	69[a]	41[b]	47[c]

Note: Different superscripts on the same line denote statistically significant differences using Tukey's multiple comparisons tests ($p < .05$).

A 3 (instruction groups) X 2 (gender) X 13 (scales) X 2 (endorsement) ANOVA on mean Z-scores for response latency produced significant interactions for gender X endorsement and for instruction groups X scales X endorsement, $F(1,54) = 6.28$, $p < .05$ and $F(24,648) = 3.43$, $p < .0001$, respectively. A simple effects ANOVA of gender at levels of endorsement was significant. Males took significantly longer than females to respond to items they accepted. Table II shows the means and statistically significant differences for males and females on items that were accepted and rejected.

TABLE II

MEAN Z-SCORES FOR RESPONSE LATENCY
AT LEVELS OF GENDER AND ENDORSEMENT

	Gender		
	Males	Females	Difference
Endorsement			
Accepted	0.22	-0.11	0.33*
Rejected	0.11	-0.08	0.19

*p < .05

A simple effects ANOVA of endorsement at various levels of group and scale were significant. The fake bad group required significantly less time to accept than reject items across the F, Hs, Pd, and Pt scales, while, on the L scale they required more time to accept items. Subjects instructed to respond honestly took significantly more time to accept than reject items on the F and Sc scales, a pattern that appeared for the fake good subjects on the F scale as well. Table III shows statistically significant diffences for the simple effects ANOVA's of endorsement at levels of instruction groups X scales.

TABLE III

MEAN Z-SCORES FOR RESPONSE LATENCY AT LEVELS OF ENDORSEMENT, INSTRUCTION GROUP, AND SCALE

Group	Scale	Endorsement: Accepted	Endorsement: Rejected	Difference
Fake Bad				
	L	0.71	-0.07	0.78*
	F	0.04	0.32	0.28*
	K	0.05	-0.21	0.26
	Hs	0.04	0.69	0.64*
	D	0.10	0.11	0.01
	Hy	0.14	0.18	0.04
	Pd	-0.17	0.10	0.27*
	Mf	-0.09	0.09	0.18
	Pa	-0.17	0.09	0.26
	Pt	-0.19	0.15	0.34*
	Sc	-0.11	0.11	0.22
	Ma	-0.27	-0.09	0.18
	Si	-0.18	-0.14	0.04
Fake Good				
	L	0.17	0.29	0.08
	F	0.37	-0.12	0.49*
	K	-0.06	0.13	0.19
	Hs	-0.10	-0.01	0.09
	D	0.11	-0.01	0.12
	Hy	0.02	0.02	0.00
	Pd	0.04	-0.06	0.10
	Mf	0.06	0.12	0.06
	Pa	-0.03	-0.11	0.08
	Pt	0.02	-0.17	0.19
	Sc	0.06	-0.14	0.20
	Ma	0.10	-0.11	0.21
	Si	0.11	-0.11	0.22
Honest				
	L	0.15	-0.06	0.21
	F	0.43	-0.10	0.53*
	K	-0.01	0.04	0.05
	Hs	0.02	-0.10	0.12
	D	0.02	-0.02	0.04
	Hy	0.02	-0.10	0.12
	Pd	0.17	-0.07	0.24

Scale	Accepted	Rejected	Difference
Mf	0.12	0.02	0.10
Pa	0.10	-0.02	0.12
Pt	0.07	-0.02	0.09
Sc	0.35	-0.09	0.44*
Ma	0.04	-0.07	0.11
Si	0.02	0.05	0.03

*$p < .05$

Simple effects ANOVA's on instruction groups at various levels of endorsement and scale, as well as Tukey's multiple comparisons test revealed a few significant differences. Fake bad subjects took more time to accept items on the L scale than either fake good or honest subjects, which did not differ from one another. On the Hs scale fake bad subjects produced longer latencies when rejecting items than fake good or honest subjects. All other simple effects comparisons among groups at levels of endorsement and scales were not statistically significant.

The second overall ANOVA, a 3 (instruction groups) X 2 (gender) X 5 (scales) X 2 (endorsement) X 2 (item subtlety), produced significant interactions for instruction groups X endorsement X item subtlety and for instruction groups X scales X item subtlety, $F(2,54) = 11.64$, $p < .0001$ and $F(8,216) = 2.86$, $p < .01$, respectively. Simple effects ANOVA's for item subtlety at various levels of group and endorsement revealed some significant differences for rejected items only. The fake bad subjects required significantly more time to reject obvious items than subtle items. The opposite pattern was indicated for fake good

subjects, who took less time to reject obvious items than subtle items. No differences in rejecting subtle/obvious items appeared for the honest subjects. Table IV illustrates the differences between subtle and obvious items at levels of group and endorsement.

TABLE IV

MEAN Z-SCORES FOR RESPONSE LATENCY AT LEVELS OF INSTRUCTION GROUP, ITEM SUBTLETY, AND ENDORSEMENT

		Item Subtlety		
Group	Endorsement	Obvious	Subtle	Difference
Fake Bad				
	Accepted	-0.08	-0.08	0.00
	Rejected	0.38	-0.10	0.48*
Group	Endorsement	Obvious	Subtle	Difference
Fake Good				
	Accepted	0.18	0.02	0.16
	Rejected	-0.12	0.19	0.31*
Honest				
	Accepted	0.15	0.03	0.12
	Rejected	-0.08	0.02	0.10

*$p < .05$

Simple effects ANOVA's and Tukey's tests for instruction groups at levels of item subtlety and endorsement revealed some significant differences. Fake bad subjects took significantly longer to reject obvious items than the fake good subjects. The simple effects ANOVA's suggested that neither faking group produced statistically

significant differences when compared to the honest group.

Simple effects ANOVA's and Tukey's tests on item subtlety at levels of instruction groups and scales revealed some significant differences. Fake bad subjects took significantly longer to respond to obvious items than subtle items across the D, Hy, and Pa scales. Table V shows significant differences for the fake bad group as well as nonsignificant results for the fake good and honest groups. Neither the honest nor fake good groups produced differences across subtle/obvious scales. In addition, no significant differences were observed in simple effects ANOVA's on instruction groups at levels of item subtlety and scale.

TABLE V

MEAN Z-SCORES FOR RESPONSE LATENCY AT LEVELS OF INSTRUCTION GROUP, SCALE, AND ITEM SUBTLETY

		Item Subtlety		
Group	Scale	Obvious	Subtle	Difference
Fake Bad				
	D	0.24	-0.11	0.35*
	Hy	0.38	0.10	0.48*
	Pd	0.12	-0.06	0.18
	Pa	0.19	-0.18	0.37*
	Ma	-0.17	-0.18	0.01
Fake Good				
	D	0.02	0.22	0.20
	Hy	-0.04	0.16	0.20
	Pd	0.24	0.04	0.20
	Pa	-0.12	0.12	0.24
	Ma	0.04	-0.02	0.06

Honest				
	D	0.00	0.08	0.08
	Hy	0.08	-0.07	0.15
	Pd	0.12	0.04	0.08
	Pa	0.04	0.04	0.00
	Ma	-0.06	0.02	0.08

*$p < .05$

CHAPTER VI

DISCUSSION

Current findings

The initial ANOVA on T-scores for validity and clinical scales indicated that subjects responded differentially to the instructions. Subjects instructed to fake bad produced the most dramatic effects, scoring significantly higher than subjects instructed to fake good or respond honestly on all the clinical scales and the F scale. Fake bad subjects also scored significantly lower than the other two groups on the L and K scales. Instructions to fake good appeared to be less effective. Subjects faking good scored higher than subjects instructed to respond honestly on the L and K scales, and lower on the F and Si scales. All the significant differences for the fake bad and fake good groups were consistent with the expected effects of faking on the MMPI-2. However, the number of differences between subjects instructed to fake good and respond truthfully was considerably lower than for the fake bad group. The problem may be related to the "floor effect" described by Peterson, Clark, and Bennett (1989). Those researchers stated that honestly responding college students endorsed so few psychopathology-related items that when asked to fake good few differences were noted.

The results of statistical analysis indicated that MMPI-2 response latencies were affected by instructions to

fake, particularly instructions to simulate psychopathology. In general, subjects required more time to respond to an item when the item was contrary to the response set the subject was operating under. The strategy of examining item endorsement was particularly helpful in grouping items for analysis.

In the first three-way interaction of instruction groups X scales X endorsement subjects instructed to fake bad required significantly less time to accept items from the F, Hs, Pd, and Pt scales than to reject those items. These scales are composed of many or almost entirely of items which are psychopathology-related. The fake bad subjects, however, took longer to accept items on the L scale, which is composed of positive or psychopathology-unrelated items. The opposite pattern of results was true of the honest and fake good subjects. Subjects instructed to repond honestly took more time to accept items on the F and Sc scales, while fake good subjects took longer to accept items on the F scale. Both the F and Sc scales are psychopathology-related.

Markus (1977) and Kuiper (1981) both hypothesized that a self-schema would facilitate processing of schema-related material. Those researchers demonstrated that items which were consistent with the self-schema or response set the subject was operating under were answered more quickly. For fake bad subjects in the current study, items related to a psychopathology schema (F, Hs, Pd, and Pt scales) were more

quickly accepted, items not related to a psychopatholgy schema (L scale) were more quickly rejected. For the honest and fake good subjects, items related to psychopathology (F and Sc scales) were more quickly rejected. The response set or schema appeared to have facilitated processing of schema-relevant information and hindered processing of schema-unrelated information. This finding is directly contradictory to the theoretical position of Rogers (e.g., Rogers, Kuiper, & Kirker). Rogers postulated that items which were representative of an individual would take longer to process and thus would have predicted that MMPI-2 items relevant to the instruction set would be more slowly accepted.

Examination of the second three-way interaction of instruction groups X accepted/rejected items X subtle/obvious subscales interaction revealed significant differences for rejected items only. The fake bad subjects required significantly more time to reject obvious items than subtle ones. The opposite pattern was indicated for fake good subjects, who took less time to reject obvious items than subtle items. No differences in rejecting subtle/obvious items appeared for the subjects instructed to respond honestly. When comparing among groups, the fake bad subjects took significantly longer to reject obvious items than the fake good subjects. Neither faking group differed from the honest group on rejecting subtle/obvious items.

The schema theory of Markus (1977) and Kuiper (1981) appeared to provide the best explanation for subjects rejection of subtle and obvious items. Christian, Burkhart, and Gynther (1978) reported a significant negative correlation (r = -.78) between rated obviousness of items and social desirability. Given the correlation, the obvious subscales represented psychopathology-related items and the subtle subscales represented psychopathology-unrelated items. The self-schema theory appears to offer the best interpretation of the data. Using the schema theory of Markus (1977) and Kuiper (1981), the increased latencies of fake bad subjects in rejecting obvious items indicated that such a decision is counter to the schema the subjects were operating under and, therefore, increased the processing time. For the fake good subjects the opposite pattern was observed, where latencies were decreased in rejecting obvious items. Rejecting a psychopathology-related item would be consistent with a fake good or honest schema and response time would be shorter than for psychopathology-related items. When comparing fake bad to fake good subjects, fake bad would be expected to take longer in rejecting obvious items; that hypothesis was confirmed in the data.

In addition to Kuiper (1981) and Markus (1977), the findings of Popham and Holden (1990) were supported. For fake bad subjects, scaled scores were expected to be higher than honest or fake good. Popham and Holden (1990)

demonstrated that if a person scored high on a scale, items unrepresentative of that person (rejected items) would have longer latencies than representative items (accepted items). Fake bad subjects quickly accepted items on psychopathology scales (F, Hs, Pd, Pt) and took longer to accept items on a non-psychopathology scale (L scale). Fake good and honest subjects, on the other hand, took longer to accept items on pathology scales (F and Sc) and less time to reject psychopathology items (obvious subscales).

Examination of the instruction groups X scales X item subtlety interaction produced confusing results. When collapsing across endorsement, fake bad subjects took significantly longer to respond to obvious items than subtle items across the D, Hy, and Pa scales. Those findings are contradictory to the above mentioned results. Using the schema theory of Kuiper (1981) and Markus (1977), fake bad subjects would be expected to respond quickly to obvious items because of their obvious relationship to psychopathology. Subtle items should take longer to respond to because they are unrelated to a psychopathology schema. No explanation consistent with the earlier findings can be offered, although the contradictory data can be explained in isolation.

The notion of social desirability threshold from Rogers (1971) offers some explanation of the conflicting data. Two assumptions are necessary to apply the Rogers (1971) theory. First, it must be assumed that the Wiener/Harmon subtle

items are most closely representative of normal college students. These are items which by definition are not clearly related to psychopathology. Second, we must assume that a fake bad instruction set "shifts" the social desirability threshold from subtle (psychopathology-unrelated) to obvious (psychopathology-related) items. In other words, the threshold shift in subjects instructed to fake bad makes them more willing to endorse socially undesirable items. Items near the threshold have a nearly equal chance of being answered true or false, they are "controversial" items, and are more difficult for subjects to answer. Therefore, according to the Rogers (1971) theory, the fake bad subjects required more time to answer obvious items because the instruction set shifted their threshold into the obvious, psychopathology-related items. This explanation, however, is in direct contradiction to the theories of Markus (1977) and Kuiper (1981).

Only three main effects for instruction group were significant. When rejecting obvious items, rejecting items on the Hs scale, and accepting items on the L scale, fake bad subjects required more time than fake good subjects or subjects instructed to respond honestly. These differences again support the schema theory (e.g., Markus, 1977) prediction that schema or response set relevant items will be more quickly processed. However, statistically significant differences were almost entirely limited to patterns of responses within instruction groups. The

extrapolated theories of Rogers (e.g., Rogers, Kuiper, & Kirker, 1977), Nowakowska (1970), and Kuncel (1973), which predicted among group differences, were not supported by the results. The problem may be with the assumption that faking involves accessing a less elaborate information network or self-schema. In addition, no significant differences in response latencies occurred between the fake good and honest groups. Both groups appeared to process the MMPI-2 items with approximately the same speed. The similarity of response time for subjects instructed to respond honestly and to fake good may be a demonstration of the the "floor effect" described by Peterson et al. (1989).

Follow-up tests on the gender X endorsement interaction indicated that males took significantly longer than females to accept all items. Further investigations into the connection between MMPI-2 response latencies and faking must take gender differences in decision speed into account. In addition, it may be necessary to develop separate norms for males and females with future empirical studies.

Future Research

When comparing the patterns of responses within the fake bad and fake good/honest groups it was clear that the items were perceived and processed differently. Significant differences were mainly observed within-groups of instructed subjects. In general, subjects accepted and rejected items consistent with their response set faster than items

inconsistent with their response set. Further research must focus on the inidividual items that differentiate the groups. In much the same way that Hathaway and McKinley (e.g., 1942) developed the original MMPI scales, patterns of differences on latencies could be used to identify items which are suseptible to over- or under-reporting of psychopathology. The differences in response latency patterns may be applicable in distinguishing between honest and faking examinees, particularly when over-reporting of psychopathology is suspected.

Further study of response latencies and faking on the MMPI-2 is needed, given the ease with which latencies can be gathered when the MMPI-2 is computer-administered and the possibility of their use as validity measures. Researchers should consider using the Z-score standardization procedure described in this study. In addition, any examination of MMPI-2 latencies must involve subtle/obvious subscales and accepted/rejected items. Future refinements might include the use of clinical populations with reasons to distort their responses, given the difference in effectiveness of faking instructions in this study as evidenced by considerably fewer differences in scale T-scores. Use of clinical subjects could avoid the "floor effect" demonstrated in this study and described by Peterson et al. (1989). Use of the MMPI-2 content scales may also be helpful. If subjects are instructed to fake specific psychological problems, such as anxiety, and corresponding

MMPI-2 content scales are examined (i.e., ANX), response latencies may be more clearly related to faking.

REFERENCES

Anthony, N. (1971). Comparison of clients' standard, exaggerated, and matching MMPI profiles. Journal of Consulting and Clinical Psychology, 36, 100-103.

Best, J. B. (1986). Cognitive Psychology. St. Paul, MN: West Publishing.

Burkhart, B. R., Christian, W. L., & Gynther, M. D. (1978). Item subtlety and faking on the MMPI: A paradoxical relationship. Journal of Personality Assessment, 42, 76-80.

Butcher, J. N. (1990). MMPI-2 in Psychological Treatment. New York: Oxford.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1990). Minnesota Multiphasic Personality Inventory (MMPI-2). Manual for administration and scoring. Minneapolis: University of Minnesota Press.

Cantor, N., & Mischel, W. (1977). Traits as prototypes: Effects on recognition memory. Journal of Personality and Social Psychology, 35, 38-48.

Christian, W. L., Burkhart, B. R., & Gynther, M. D. (1978). Subtle-obvious ratings of MMPI items: New interest in an old concept. Journal of Consulting and Clinical Psychology, 46, 1178-1186.

Costa, P. T., Zonderman, A. B., Williams, R. B., & McCrae, R. R. (1985). Content and comprehensiveness in the MMPI: An item factor analysis in a normal adult sample. Journal of Personality and Social Psychology, 48, 925-933.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General, 104, 268-294.

Dahlstrom, W. G., Welsh, G. S., & Dahlstrom, L. E. (1972). An MMPI handbook: Vol. 1. Clinical Interpretation. Minneapolis, MN: University of Minnesota Press.

Dubinsky, S., Gamble, D. J., & Rogers, M. L. (1985). A literature review of subtle-obvious items on the MMPI. Journal of Personality Assessment, 49, 62-68.

Duff, F. L. (1965). Item subtlety in personality inventory scales. Journal of Consulting Psychology, 29, 565-570.

Dunn, T. G., Lushene, R. E., & O'Neil, H. F., Jr. (1972). Complete automation of the MMPI and a study of its response latencies. Journal of Consulting and Clinical Psychology, 39, 381-387.

Gough, H. G. (1947). Simulated patterns on the Minnesota Multiphasic Personality Inventory. Journal of Abnormal and Social Psychology, 42, 215-225.

Gough, H. G. (1950). The F minus K dissimulation index for the MMPI. Journal of Consulting Psychology, 14, 408-413.

Graham, J. R. (1987). The MMPI: A practical guide (2nd ed.). New York: Oxford.

Graham, J. R., & Lilly, R. S. (1984). Psychological testing. Englewood Cliffs, NJ: Prentice-Hall.

Grayson, H. M. (1951). A psychological admissions testing program and manual. Los Angeles: Veterans Administration Center, Neuropsychiatric Hospital.

Greene, R. L. (1980). The MMPI: An interpretive manual. New York: Grune & Stratton.

Greene, R. L. (1988). The relative efficacy of F - K and the obvious and subtle scales to detect overreporting of psychopathology on the MMPI. Journal of Clinical Psychology, 44, 152-159.

Groth-Marnat, G. (1990). Handbook of psychological assessment (2nd ed.). New York: John Wiley & Sons.

Grow, R., McVaugh, W., & Eno T. D. (1980). Faking and the MMPI. Journal of Clinical Psychology, 36, 910-917.

Hanley, C. (1962). The "difficulty" of personality inventory items. Educational and Psychological Measurement, 22, 577-584.

Harvey, M. A., & Sipprelle, C. N. (1976). Demand characteristic effect on the subtle and obvious subscales of the MMPI. Journal of Personality Assessment, 40, 539-544.

Hathaway, S. R., & McKinley, J. C. (1940). A multiphasic personality schedule (Minnesota): I. Construction of the schedule. Journal of Psychology, 10, 249-254.

Hathaway, S. R., & McKinley, J. C. (1942). A multiphasic personality schedule (Minnesota): III. The measurement of symptomatic depression. Journal of Psychology, 14, 73-84.

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

Hathaway, S. R., & McKinley, J. C. (1967). The Minnesota Multiphasic Personality Inventory (revised manual). New York: Psychological Corporation.

Honaker, L. M. (1988). The equivalency of computerized and conventional MMPI administration: A critical review. Clinical Psychology Review, 8, 561-577.

Hsu, L. M., Santelli, J., & Hsu, J. R. (1989). Faking detection validity and incremental validity of response latencies to MMPI subtle and obvious items. Journal of Personality Assessment, 53, 278-295.

Jackson, D. N. (1967). Manual for the Personality Research Form. Goshen, NY: Research Psychologists Press.

Jacoby, L. L., & Craik, F. I. M. (1979). Effects of elaboration of processing at encoding and retrieval: Trace distinctiveness and recovery of initial context. In L. S. Cermack & F. I. M. Craik (Eds.), Levels of processing in human memory. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kuiper, N. A. (1981). Convergent evidence for the self as a prototype: The "inverted-U RT effect" for self and other judgements. Personality and Social Psychology Bulletin, 7, 438-443.

Kuiper, N. A., & Rogers, T. B. (1979). Encoding of personal information: Self-other differences. Journal of Personality and Social Psychology, 37, 499-514.

Kuncel, R. B. (1973). Response processes and relative location of subject and item. Educational and Psychological Measurement, 33, 545-563.

Markus, H. (1977). Self-schemata and processing information about the self. Journal of Personality and Social Psychology, 35, 63-78.

Meehl, P. E. (1945). The dynamics of "structured" personality tests. Journal of Clinical Psychology, 1, 296-303.

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

Nowakowska, M. (1970). A model of answering to a questionnaire item. Acta Psychologica, 34, 420-439.

Peterson, G. W., Clark, D. A., & Bennett, B. (1989). The utility of MMPI subtle, obvious scales for detecting fake good and fake bad response sets. Journal of Clinical Psychology, 45, 575-582.

Popham, S. M., & Holden, R. R. (1990). Assessing MMPI constructs through the measurement of response latencies. Journal of Personality Assessment, 54, 469-478.

Rogers, T. B. (1971). The process of responding to personality items: Some issues, a theory, and some research. Multivariate Behavioral Research Monographs, 6(2), 1-65.

Rogers, T. B. (1974a). An analysis of the stages underlying the process of responding to personality items. Acta Psychologica, 38, 205-213.

Rogers, T. B. (1974b). An analysis of two central stages underlying responding to personality items: The self-referent decision and response selection. Journal of Research in Personality, 8, 128-138.

Rogers, T. B. (1977). Self-reference in memory: Recognition of personality items. Journal of Research in Personality, 11, 295-305.

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. Journal of Personality and Social Psychology, 35, 677-688.

Seeman, W. (1952). "Subtlety" in structured personality tests. Journal of Clinical Psychology, 16, 278-283.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder's methods. Acta Psychologica, 30, 276-315.

Tarter, R. E. (1988). A review of the Minnesota Multiphasic Personality Inventory. In M. Herson and A. Bellack (Eds.), Dictionary of Assessment Techniques, New York; Pergamon Press.

Temple, D. E., & Geisinger, K. F. (1990). Response latency to computer-administered inventory items as an indicator of emotional arousal. Journal of Personality Assessment, 54, 289-297.

Vesprani, G. J., & Seeman, W. (1974). MMPI X and zero items in a psychiatric outpatient group. Journal of Personality Assessment, 38, 61-64

Wales, B., & Seeman, W. (1968). A new method for detecting the fake-good response set on the MMPI. Journal of Clinical Psychology, 24, 211-216.

Wales, B., & Seeman, W. (1969). What do MMPI zero items really measure: An experimental investigation. Journal of Clinical Psychology, 25, 420-424.

Wales, B., & Seeman, W. (1972). Instructional sets and MMPI items. Journal of Personality Assessment, 36, 164-170.

Wiener, D. N. (1948). Subtle and obvious keys for the MMPI. Journal of Consulting Psychology, 12, 164-170.

Wiener, D. N., & Harmon, L. R. (1946). Subtle and obvious keys for the MMPI: Their development. (Advisement Bulletin No. 16, Regional Veterans Administration Office, Minneapolis).

Worthington, D. L., & Schlottmann, R. S. (1986). The predictive validity of subtle and obvious empirically derived psychological test items under faking conditions. Journal of Personality Assessment, 50, 171-181.

APPENDIX A

APPROVAL TO USE COPYRIGHTED MMPI-2 ITEMS

UNIVERSITY OF MINNESOTA

Twin Cities Campus *University of Minnesota Press* *2037 University Avenue S.E.*
Minneapolis, MN 55414-3092
612-624-2516
612-624-0005
1-800-388-3863 (orders only)
Fax: 612-626-7313

September 21, 1992

David Brunetti
215 N. Murray Hall
Oklahoma State University
Stillwater, OK 74078

David Brunetti:

As requested in your letter of September 9, we are extending until July 1, 1993, your permission to reproduce the first 370 items of the MMPI-2 test booklet on computer screen for the purpose of collecting data for your dissertation. A reminder that we require that you destroy all computer files and disks containing MMPI items when your research is complete.

Cordially,

Beverly Kaemmer
MMPI Manager

APPENDIX B

INFORMED CONSENT STATEMENT

INFORMED CONSENT STATEMENT

Project title: Response styles on the Minnesota Multiphasic Personality Inventory - 2.
Experimenters: David Brunetti, M.S. & Robert Schlottmann, Ph.D.

I, (print name) _____ hereby authorize and direct David Brunetti and Robert Schlottmann, or assistants of their choosing, to perform the procedures listed here.

A. Purpose: This study is designed to examine different response styles on the Minnesota Multiphasic Personality Inventory - 2 (MMPI-2). The MMPI-2 is a widely-used, standardized personality test. The format of the MMPI-2 consists of several hundred true/false items, some of which may be personal in nature.

B. Procedures: You will be asked to do the following:

1. Read a set of instructions on using the computer and responding to the MMPI-2.
2. Answer items on a shortened form of the MMPI-2.
3. Participate in a debriefing session at the end of the experiment. Your questions will be addressed at that time. Please note that no information gained from the MMPI-2 will be made available to you.

C. Duration: Your participation will take approximately one hour.

D. Confidentiality: Computer files of this experiment's data will be numerically coded. You will not be affixing your name to any of the MMPI-2 materials. Thus, you anonymity will be assured.

E. Risks: The risks in this study are minimal and do not exceed those ordinarily encountered in daily life. Some people may find specific items on the MMPI-2 personal and/or intrusive, but they are part of routine psychological evaluation and testing.

F. Benefits: Through your participation in this study you will be exposed to scientific psychological research. This may help you in understanding the procedures and methods of psychology. In addition, the results of this and subsequent related studies may aid psychologists in understanding responses on the MMPI-2. You will also be awarded extra credit points for your participation.

G. Compensation for participation: You will be awarded 1 extra credit point in your PSYCH 1113 class for each hour or fraction of an hour you choose to participate in this

experiment; there are other ways that you can get extra credit in your class. You can be involved in other experiments and/or complete projects (e.g., book reports) which your instructor will explain.

I have been fully informed about the procedures listed here.

I am aware of what I will be asked to do and of the risks and benefits in this study. I also understand the following statements:

I certify that I am at least 18 years of age.

My participation today is part of an investigation called Response Styles on the Minnesota Multiphasic Personality Inventory - 2.

The purpose of this study is to examine different response styles on the MMPI-2.

I understand that my participation is voluntary, that there is no penalty for refusal to participate, and that I am free to withdraw my consent and participation at any time and without penalty.

I understand that I may contact any of the experimenters at the following address and telephone number should I desire to discuss my participation in this study and/or to request information about the study's outcome: 215 North Murray Hall, Department of Psychology, Oklahoma State University, Stillwater, OK 74078-0250, 405/744-6027. Additionally, I understand that I may contact the Office of University Research Services, 001 Life Sciences East, Oklahoma State University, Stillwater, OK 74078, 405/744-9991.

I have fully read and understand this consent form. I sign if freely and voluntarily. A copy of this form has been given to me. I hereby give permission for my participation.

_____ _____

Signature of participant Date

_____

Signature of witness

APPENDIX C

COMPUTER INSTRUCTIONS

In a moment you will take the Minnesota Multiphasic Personality Inventory -2 (MMPI-2). This inventory consists of numbered statements. You are to read each statement and decide if it is TRUE (or MOSTLY TRUE) or FALSE (or MOSTLY FALSE). You are to decide on your own; the experimenter will provide no assistance or clarification. After deciding, you will answer on the computer keyboard in front of you. If an item is TRUE (or MOSTLY TRUE), you will press the 1 key located above the Q key. If an item is FALSE (or MOSTLY FALSE), you will press the 2 key located above the letter W. You will not be using the numeric keypad on the right side of you keyboard.

You must respond with either a 1 or a 2, even if you cannot decide. Also, you will not be allowed to review your answers or go back to change them.

Each MMPI-2 item will be presented on a screen similar to the following. After you have responded either 1 or 2, the screen will go blank and you will be asked to press the space bar to display to next item.

Item Number 1

I enjoy walking my dog.

1 True 2 False

VITA

David Gene Brunetti

Candidate for the Degree of

Doctor of Philosophy

Thesis: THE EFFECTS OF FAKING INSTRUCTIONS ON MMPI-2 RESPONSE LATENCIES

Major Field: Clinical Psychology

Biographical:

Personal Data: Born in Berkeley, California, On December 22, 1964, the son of Lawrence and Virginia Brunetti.

Education: Graduated from Rampart High School in Colorado Springs, Colorado in May 1983; received a Bachelor of Science degree in Psychology from Colorado State University, Fort Collins, Colorado in December 1987. Completed the requirements for the Master of Science degree in Psychology at Oklahoma State University, Stillwater, Oklahoma in December 1989. Doctor of Philosophy in Clinical Psychology to be conferred in December 1994.

Experience: Undergraduate research assistant in animal and human psychological studies at Colorado State University; Research assistant, graduate teaching assistant, and clinical experience through Oklahoma State University, Department of Psychology, 1988 to 1993; Clinical Psychology internship at the Missouri Health Sciences Psychology Consortium, Columbia, Missouri, 1993 to 1994.

Professional Memberships: American Psychological Association, Association for Advancement of Behavior Therapy, Society for Personality Assessment.

OKLAHOMA STATE UNIVERSITY
INSTITUTIONAL REVIEW BOARD
FOR HUMAN SUBJECTS RESEARCH

Proposal Title: RESPONSE STYLES ON THE MMPI-2

Principal Investigator: Dr. Robert Schlottmann/David Brunetti

Date: June 16, 1992 IRB # AS-92-054

This application has been reviewed by the IRB and

Processed as: Exempt [ ] Expedite [XX] Full Board Review [ ]

Renewal or Continuation [ ]

Approval Status Recommended by Reviewer(s):

Approved [XX] Deferred for Revision [ ]

Approved with Provision [ ] Disapproved [ ]

Approval status subject to review by full Institutional Review Board at next meeting, 2nd and 4th Thursday of each month.

Comments, Modifications/Conditions for Approval or Reason for Deferral or Disapproval:

Comments:
Regarding the consent form, I would offer a suggestion of Item E. The researcher might consider adding a reiteration that the participant's name will not appear on the answer sheet; therefore, there will be no means to detect individual responses to items which the person may consider to be intrusive or personal.

It is stated that participants will receive one extra point of credit for each hour and each fraction of hour of time spent responding. Is it possible that some will deliberately take more than the hour predicted in order to earn the extra point and if they do, will this have any impact upon the results obtained?

Signature: Marcia K. Tilley Date: July 10, 1992
Chair of Institutional Review Board