

A DISTINCTION BETWEEN ZERO-X AND SUBTLE-OBVIOUS
DIMENSIONS OF MMPI ITEMS AND THE VARIANCE
OF MMPI SCORES UNDER DIFFERENT
INSTRUCTIONAL SETS

A THESIS
SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF ARTS
IN THE GRADUATE SCHOOL OF THE
TEXAS WOMAN'S UNIVERSITY

COLLEGE OF EDUCATION

BY
JOAN MARY THOMPSON, B.S.

DENTON, TEXAS

AUGUST 1981

The Graduate School
Texas Woman's University
Denton, Texas

_____ August 17 _____ 19 81

We hereby recommend that the _____ thesis _____ prepared under
our supervision by _____ Joan Mary Thompson _____

Dissertation/Theses signature page is here. To protect individuals we have covered their signatures.

Thesis
T1981
T453d
c.2

TEXAS WOMAN'S UNIVERSITY LIBRARY

TABLE OF CONTENTS

LIST OF TABLES	iv
LIST OF FIGURES	vi
INTRODUCTION	1
Item Validity	2
Dimensions of Content	2
Subtle-Obvious	3
Zero-X	10
Sources of Variance	12
Test Attitude	12
Variance	15
Response Sets	17
Demand Characteristics	23
Impression Management	24
Test-Wiseness	25
Faking Ability	26
Effects of Faking on Variance	28
Motivation	29
Cognitive vs. Affective Responding	30
METHOD	32
Subjects	32
Procedure	32
RESULTS	35
DISCUSSION	54
Instructional Sets	54
Zero-X Items	58
APPENDIX A	65
REFERENCES	68

5048

82-Direct

LIST OF TABLES

1.	Scale Distribution of Zero and X Items by Degree of Subtlety	13
2.	Means and Standard Deviations for Proportions of Zero-X Items Endorsed Within Each Sub- tlety Category for Instructional Set Groups . . .	36
3.	Means and Standard Deviations for Proportions of Zero-X Items Endorsed Within Each Sub- tlety Category for Groups Formed According to F-K Scores	38
4.	Mean Differences and t Values on Zero-X Items for Fake Bad, Look Bad, Look Good, and Fake Good Profiles	41
5.	Analysis of Variance Summaries for Three Regression Equations with the K Scale, the F Scale, and the F-K Index as Cri- terion Measures	42
6.	Significance Tests of Regression Coefficients with the F Scale as the Criterion	43
7.	Significance Tests of Regression Coefficients with the K Scale as the Criterion	45
8.	Significance Tests of Regression Coefficients with the F-K Index as the Criterion	46

9.	Beta Weights and Multiple Correlation	
	Coefficients for Three Criterion Variables . . .	47
10.	Means and Standard Deviations of MMPI Scales	
	on Fake Bad, Look Bad, Look Good, and	
	Fake Good Profiles	49
11.	Levene's Test for Overall Equality of Variances	
	of MMPI Scales and Pairwise Variance Compari-	
	sons Between Fake Bad Profiles and Look Bad,	
	Look Good, and Fake Good Profiles,	
	Respectively	51
12.	Zero-X and Subtle-Obvious Classifications	
	for Double-Scored MMPI Items	62

LIST OF FIGURES

1. Mean scores on ZVS, ZSS, ZN, XVS, XSS,
and XN items for Fake Bad, Look Bad,
Look Good, and Fake Good groups 39

A Distinction Between Zero-X and Subtle-Obvious
Dimensions of MMPI Items and the Variance
of MMPI Scores Under Different
Instructional Sets

In the thirty-eight years since it was published, the Minnesota Multiphasic Personality Inventory (MMPI) has become the most widely researched and written about instrument in the history of testing (Buros, 1978, p. xxxviii), as well as the instrument of choice in a clinical setting when structured personality tests are used. In a review for The Eighth Mental Measurements Yearbook, Henry Alker wrote,

The MMPI is an old soldier with some 6,000 citations on its chest. Quite a few of these decorations, to be sure, are combat ribbons commemorating campaigns in which the MMPI lost rather than won the important battles. And many more of these ribbons were earned by merely participating in the struggle without materially affecting the outcome one way or another. (Buros, 1978, p. 931)

Much of the controversy over the MMPI has centered on the question of validity. Some of its critics believe that low validity is inherent in the group-discriminative method of test construction which can result in items of ambiguous

value (Jackson, 1971). Others argue that the structure of the test itself causes responses to be biased (Berg, 1957). Furthermore, many investigators appear to interpret MMPI responding in terms of evaluative judgments on the part of subjects and conduct their research accordingly. Literature relevant to item validity will be reviewed, followed by a discussion of the literature pertaining to the variance accounted for by the subject's test attitude.

Item Validity

Dimensions of Content

The MMPI contains many items whose contribution to the validity of the instrument has been questioned because their contents do not clearly reflect the expected behaviors or attitudes symptomatic of psychopathology (Jackson, 1971). Whereas pathological depression is evident when the "obvious" statement "Most of the time I wish I were dead" is answered true, there is no rational basis for explaining why "The man who provides temptation by leaving valuable property unprotected is about as much to blame for its theft as the one who steals it" is indicative of pathology, or why "I do not often notice my ears ringing or buzzing," answered false, is a sign of paranoia. The inclusion of these "subtle" items in the final inventory is the result of the criterion group-discriminative method of test construction (Cronbach, 1970, chap. 16). In the empirical approach, few a priori assump-

tions are made about the validity of individual items or their scale membership: If there is a significant difference between the means of the normal group and the criterion group, the item is considered ipso facto a valid measure of the criterion. An item with a mean difference of two or more standard deviations for the two groups was considered discriminative of the criterion (McKinley & Hathaway, 1940). Approximately 150 of the items on the validity and clinical scales are subtle in meaning.

An overlapping set of items, most of which are subtle, are known as "zero" items. Zero items are those which were endorsed by a majority of normals but are scored for pathology because they were endorsed by a greater majority of clinical cases (Meehl & Hathaway, 1946). There are 82 zero items, and they can contribute a possible 116 points to the raw score (Dahlstrom & Welsh, 1960, pp. 400-406). Items scored for pathology which were endorsed by a minority of the normal sample are known as "X" items. The pathological content in most of them is apparent. Both types of items, however, were considered pathological in the construction of the test: the obvious ones indicated the more severe forms of illness, whereas subtle ones denoted subtle dimensions (Wiener, 1948b).

Subtle-Obvious

Meehl (1945b) believed that the concept of subtlety is

inherent in structured personality tests, allowing subjects to reveal characteristics they are unable or unwilling to verbalize. Noting that some hospitalized psychiatric patients had T-scores lower than those of matched normals, Meehl (1945a) used subtle items in an experimental scale to investigate Rosanoff's theory of a "control" factor present in subjects who are in fact abnormal, yet are able to function effectively. His research gave no support to the hypothesis, but it led to the development of the K scale, a measure intended to detect either those who deny (dissimulators) or exaggerate (malingerers) psychopathology (McKinley, Hathaway, & Meehl, 1948; Meehl & Hathaway, 1946). Comprised of 30 subtle items, 21 of which are also zero items, the effectiveness of K is contingent upon the inability to discern meaning in subtlety.

Harmon and Wiener (Wiener, 1948b) developed subtle-obvious (S-O) keys for five of the clinical scales: D (Depression), Hy (Hysteria), Pd (Psychopathic Deviate), Pa (Paranoia), and Ma (Hypomania). They validated them on the response frequencies of 100 males in the original Minnesota normative group. In these five scales, there were 146 obvious items and 110 subtle ones; 62 of these items were zero items. Wiener thought the subtle keys could be used to measure personality characteristics of normal populations. Scoring the protocols of students and subjects in on-the-job

training with the S-O keys, he was able to differentiate between the successful and unsuccessful ones, whereas he could not when the standard keys were used. Successful subjects had higher S scores than O scores; the reverse was true for unsuccessful subjects. The average correlation between S and O keys was $-.15$. S and O scores were also found to discriminate between successful and unsuccessful salesman (Wiener, 1948a). Wiener's findings that well-adjusted and successful individuals obtained more abnormal scores on the subtle scales than maladjusted and unsuccessful persons were significant because they were contrary to the original validation studies for the MMPI.

After factor analyzing the Harmon and Wiener subscales, Lingoes (1960) challenged Wiener's two-factor theory and concluded that there are many dimensions to both subtleness and obviousness. The point has been made that both content and scale membership should be determinants in classifying items as subtle or obvious; i.e., an item might be considered subtle on one scale and obvious on another (Edwards, 1953; Hanley, 1961).

Subtle items were considered valuable because they appeared to be impervious to detection. Seeman (1952) found that graduate and advanced undergraduate students in clinical psychology who had taken two courses in abnormal psychology were unable to determine the scale membership of selected

subtle items, nor could they identify them as being indicative of psychopathology. Upon retesting, after having taken two additional abnormal psychology courses, they were still unable to do so (Seeman, 1953).

In an investigation of the S-O dimension in which three college groups were tested under standard, fake good, and fake bad instructions, respectively, it was shown that obvious items could be manipulated; i.e., faked, to a significantly greater extent than subtle ones (Cofer, Chance, & Judson, 1949). These results were corroborated by Gloye and Zimmerman (1967) when they instructed students to respond according to real-self and then retested them under ideal-self instructions. After testing college students under honest conditions, Rosen (1956) had the students rate the items twice, according to degree of personal desirability and social desirability. The finding that subtle item scores increased under the latter conditions led Rosen to theorize that endorsement of subtle items was a measure of psychological integration. Fricke (1957) noted that most subtle items are scored in the negative direction, and he attributed the endorsement of subtle items by well-adjusted individuals to a response tendency to answer false.

Wiener (1951), using hospitalized and nonhospitalized psychiatric patients, supported his earlier findings that those who were better adjusted had higher subtle scores than

obvious scores. He concluded

that recognition and avoidance of behavior which is socially deviate . . . and being "adjusted" or "successful" tend to go together. Similarly, lack of sensitivity to or avoidance of unusual behavior . . . and lack of success in society apparently tend to go together. (p. 7)

Comparable results led Sacks and Kirtley (1972) to a different interpretation: Psychiatric patients who endorsed obvious items to a significantly greater extent than they did subtle ones may derive secondary gain from the admission of symptoms.

Other studies of clinical populations demonstrated that subjects failed to respond to subtlety in the predicted direction. Subtle item scores decreased in crisis situations (Payne & Wiggins, 1972) and increased when subjects were motivated to appear more stable (Wilcox & Krasnoff, 1967).

Although some researchers interpreted responses to subtle and obvious items in terms of content and concluded that the cumulative responses revealed personality characteristics, there were others who attributed changes in S-O scores under different instructional sets to be a function of demand characteristics of the testing situation (Harvey & Sipprelle, 1976), item ambiguity (Hiner, Ogren, & Baxter, 1969), acquiescence (Stone, 1965), and social desirability (Wiggins,

1959, 1966).

Investigations to assess the diagnostic utility of the MMPI showed that item validity was not invariant across populations. McCall's (1958) study of depressed patients, and Manson's (1949) work with alcoholics indicated that subtlety did not contribute to the validity of the D and Pd scales, respectively.

Research on the psychoneurotic scales (Hs/Hypochondriasis, D and Hy) produced similar results. Brozek and Erickson (1948) induced a temporary, nonspecific neurosis in a group of normals by subjecting them to experimental semistarvation. A comparison of the mean scores on the neurotic scales obtained during the control period with those obtained at the end of the semistarvation period revealed significant elevations on all three scales; however, an item analysis showed that only three of the 27 items classified as subtle changed in the expected direction. Winne (1951) used two pairs of normal-neurotic groups to validate and crossvalidate neurotic triad items. A re-analysis of Winne's data using subtle-obvious ratings developed by Christian, Burkhart, and Gynther (1978) and a cutting point of 3.00 shows that 44% of the obvious items and only 12% of the subtle ones discriminate between the normal-neurotic groups at the .01 level of significance.

Although Meehl (1945a) and Wiener (1948b) both had sug-

gested that the subtle-obvious dimension had an underlying continuum, it was treated as a dichotomy until Duff (1965), using 58 doctoral students in psychology as judges, classified 226 MMPI items as subtle, intermediate, or obvious. He analyzed the item responses on the Hy, Pd, and Sc (Schizophrenia) scales of the three corresponding criterion groups and those of the 541 normals in the revised Minnesota normative group. Over 90% of the obvious items, but only 40% of the subtle ones discriminated at the same level of significance used in the original validation studies. The percentages of items which failed to discriminate were 32, 44, and 12 for Hy, Pd, and Sc, respectively. Duff's findings lend support to those who advocate content validity in personality scale items (Goldberg & Slovic, 1967; Holden & Jackson, 1979; Norman, W. T., 1963) and seriously question the relationship between subtle items and pathology. Hamsher (1969) and Cronbach (1970, p. 532) believe this relationship should be fully explored to determine if subtlety contributes to validity.

Despite the many investigations of the S-O dimension, there were no subtlety ratings for all of the items on the MMPI until Christian et al. (1978) had 138 undergraduate students in psychology classes rate the items on a 5-point scale according to degree of psychological disturbance shown. Based on the mean ratings, each item was assigned to one of

five categories: very subtle, 1.00-1.79; somewhat subtle, 1.80-2.59; neutral, 2.60-3.39; somewhat obvious, 3.40-4.19; or very obvious, 4.20-5.00. The range was 1.26-4.78 for items appearing on the validity and clinical scales.

Studies in which the S-O keys developed by Christian et al. were used to score the protocols of subjects tested under different instructional sets demonstrated the paradox discovered in previous research using less refined measures of subtlety: subtle item scores moved in the direction of greater pathology when fake good instructions were given although obvious item scores moved in the expected direction. Responding to subtle items in the scored direction appeared to indicate adjustment rather than pathology (Burkhart, Christian, & Gynther, 1978; Burkhart, Gynther, & Christian, 1978). An investigation by Gynther, Burkhart, and Hovanitz (1979) to determine the contribution of items of different levels of subtlety on the Pd scale yielded a correlation of $-.01$ between Pd-O and Pd-S, which was similar to the $.04$ obtained by Wiener (1948b).

Zero-X

Studies testing the properties of the zero-X dimension also reversed the expectations of Hathaway and McKinley that an increase in the endorsement frequencies of both zero and X items indicated greater pathology. Wales and Seeman (1968, 1972) demonstrated that zero item scores moved in opposition

to instructional set in college populations, increasing under a fake good response set and decreasing when fake bad instructions were given. Foster (cited in Wales & Seeman, 1968), in testing college students under honest, fake good, and fake bad instructions, found that subjects were unable to "successfully" manipulate zero items, whereas they were successful in manipulating X items.

Research with abnormal populations showed that zero and X items functioned as subtle and obvious ones did when subjects were motivated to alter their profiles (Anthony, 1971; Vesprani & Seeman, 1974; Wales & Seeman, 1969) supporting the findings of the S-0 dimension.

Both Duff (1965) and Christian et al. (1978) developed their S-0 keys on a rational-intuitive basis without regard for the empirical separation of items that resulted in the zero-X distinction. However, investigations by Horlick (1955) and others (Prozek & Erickson, 1948; Cofer et al., 1949; Manson, 1949; McCall, 1958; Wiener, 1948b; Winne, 1951) have shown that zero items did not retain their statistical validity across time and populations. In fact, when the response frequencies of the revised Minnesota normative group (Dahlstrom, Welsh, & Dahlstrom, 1975, pp. 193-213) are compared with those of the original Minnesota normative group, it is noted that only 69 of the 82 items retain their zero identity.

Jackson (1971), referring to questionable items as "psychologically remote" (p. 231), contended that they were the result of item sampling and never should have been included in the inventory. In the many attempts to determine their true contribution, there are numerous references to the similarity between zero and subtle items, and to the similarity between X and obvious items; however, the categories of zero-X and subtle-obvious are not mutually exclusive. Not all zero items are subtle; there are in fact more subtle X items. Table 1 shows the scale distribution of zero and X items by degree of subtlety. Furthermore, the means of the X items on the subtle-obvious continuum are not significantly higher than those of the corresponding zero items. If the zero items do discriminate there should not be a significant difference between zero and X item scores of equal degrees of subtlety. A purpose of this study is to demonstrate that there is no distinction between zero and X items of equal degrees of subtlety as measured by endorsement frequency.

Sources of Variance

Test Attitude

Of primary concern in the studies described above is the subject's test-taking attitude. Meehl and Hathaway (1946) define

the test attitude continuum by the two opposed terms "defensiveness" and "plus getting, " not im-

Table 1
Scale Distribution of Zero and X Items
by Degree of Subtlety

Scale	X Items						Zero Items						Grand total
	VS	SS	N	SO	VO	Sub-total	VS	SS	N	SO	VO	Sub-total	
<u>L</u>	1	10	3	1	0	15	0	0	0	0	0	0	15
<u>F</u>	0	4	17	23	20	64	0	0	0	0	0	0	64
<u>K</u>	1	5	3	0	0	9	6	10	5	0	0	21	30
<u>Hs</u>	1	3	19	10	0	33	0	0	0	0	0	0	33
<u>D</u>	0	8	18	15	3	44	5	9	2	0	0	16	60
<u>Hy</u>	0	6	16	18	0	40	5	12	3	0	0	20	60
<u>Pd</u>	0	5	18	13	6	42	2	5	1	0	0	8	50
<u>Mf</u> ^a	16	8	16	3	0	43	11	6	0	0	0	17	60
<u>Pa</u>	0	2	8	11	11	32	1	6	1	0	0	8	40
<u>Pt</u>	0	1	15	26	4	46	0	1	0	1	0	2	48
<u>Sc</u>	0	0	27	33	16	76	1	0	0	1	0	2	78

<u>Ma</u>	0	10	17	7	4	38	3	5	0	0	0	8	46
<u>Si</u>	5	17	22	12	0	56	4	7	3	0	0	14	70
Total	24	79	199	172	64	538	38	61	15	2	0	116	654

^aMale.

plying anything as to the degree of conscious, deliberate deception involved in either. The corresponding extremes, where such deliberate deception seems likely, we shall refer to as "faking good" and "faking bad" respectively. (p. 533)

Too often the continuum is dichotomized to a simple fake good and bad, thereby oversimplifying the concept of test attitude which, in fact, encompasses a multiplicity of psychological orientations.

The instrument itself can be a major determinant of the subject's test attitude. Personality assessment devices traditionally have been classified as structured and unstructured. Davids (1955) subdivides the structured tests into direct and indirect. Direct tests are those in which subjects are asked to rate themselves on one or more variables. Indirect tests are those in which it is assumed that subjects are not cognizant of the variables being measured. Campbell (1957) dichotomizes his classifications into objective vs. voluntary, direct vs. indirect, and free response vs. structured. In objective tests there is a correct answer, whereas in voluntary instruments any answer is acceptable. With the addition of the third dimension, objective vs. voluntary, Campbell generates eight types of instruments and classifies the MMPI as voluntary, direct, and structured.

Variance

Fiske (1971, chap. 5), in delineating six different forms of information (modes) used to measure personality, identifies generic and specific variables that account for test variance. Each of these is comprised of a subset whose elements are not necessarily mutually exclusive; e.g., the experimenter's instructions can influence the subject's motivation. Fiske classifies the MMPI as a mode 1 device; i.e., a structured, self-description questionnaire. According to Fiske, responses to mode 1 instruments can be influenced by the stimuli of the test and/or testing situation, subject motivation, the examiner's relation to the subject, and certain factors which may be inherent in the instrument itself; e.g., response sets. The importance these variables assume is mediated by whether the testing is done in an everyday life situation, wherein the interest is in the individual protocol; or as psychological research, in which the interest is in the group response.

Wiggins (1962) offered a tripartite theory of variance in the MMPI, whose definitions are quoted here because they are well stated:

By strategic variance is meant variation in test scores that may be attributed to the overall strategy of constructing scales to discriminate between criterion groups and a normative population. . . .

By stylistic variance is meant certain character-

istic response consistencies on the part of subjects which may be shown to exist relatively independently of the test itself, but whose detection is to a large extent limited by the appropriateness of a given test as a stimulus for these consistencies to emerge. . . . Method [*italics mine*] variance is [variation in test scores] due to the idiosyncratic nature of the total item pool in regard to the proportion of true and false keyings and the distribution of the item popularity values. (pp. 224-225, 241)

Not all factors which contribute to test variance are intrinsic to this thesis. The following discussion will therefore be limited to those factors which subsume test attitude as they apply to the group situation in psychological research.

Response Sets

Cronbach's (1946) paper on response sets was the most definitive to date. A response set is defined as "the consistent tendency to select responses with some common property other than the one related to the substantive variable the test is intended to measure" (Fiske, 1971, p. 298); thus, it theoretically contributes to systematic variance. Cronbach's original paper, and a subsequent one in 1950, stimulated much research to determine the influence of response sets on the validity of the MMPI. Among the sets that have

been identified as pertinent to MMPI responding are social desirability, choosing the more socially acceptable alternative; acquiescence, the tendency to agree; extremity, endorsing the more extreme alternative; and cautiousness, failure to respond to an item. The effects of the latter set are minimized in a forced-choice situation, which is true of the MMPI. However, cautiousness was clearly operant in the original validation studies when cannot say was a valid response; as many as 29% of the subjects chose this alternative to a given item (Dahlstrom & Welsh, 1960, pp. 417-429).

Extremeness in responding is a set more relevant to the individual protocol than it is to group responding in many instances. A unipolar scale, the F scale, was designed to detect individuals who operate under this set. It is comprised of 64 items present in the clinical scales that reflect the more deviant behaviors or attitudes and which were endorsed by less than 10% of the normative sample. Subjects who obtain high F scores and have otherwise normal profiles, or those who obtain high F scores and exhibit deviance on many or all of the clinical scales--in the absence of other data confirming psychopathology--are said to be malingering (Hunt, 1973). In the above quotation from Meehl and Hathaway, test attitude is conceptualized as a continuum with malingering and dissimulating representing the extremes. Gough (1947) developed the F-K dissimulation index (raw

score on F minus raw score on K) as an additional measure to detect dissembled (i.e., simulated) profiles. Using a cutting point of ± 10 , Gough was able to discriminate malingerers and dissimulators, respectively. In a subsequent study, Gough (1950) supported his earlier findings. Perlman (cited in Hunt, 1973) randomly chose 193 protocols from the Minnesota normative group and, using Gough's dissimulation index, found that 51% of the protocols were dissembled.

Investigations relative to social desirability and acquiescence have comprised the majority of MMPI research on response sets. Numerous scales purporting to measure social desirability have been developed (Cofer et al., 1949; Crowne & Marlowe, 1960; Edwards, 1957; Hanley, 1957; Wiggins, 1959). Edwards (1961) stated that his social desirability scale "provides a measure of the tendency of subjects to give socially desirable responses in self-description under the standard instructions ordinarily used with personality inventories" (p. 353). Social desirability scales are derived from items having high communality. Items with high communality are endorsed or rejected by a majority of the normative group (Wiggins, 1962); those with low communality (high controversiality) are rejected by half of the normative group and endorsed by the other half (Hanley, 1957). Messick and Jackson (1961b) compiled mean social desirability ratings for all items when keyed true; these ratings corre-

late $-.78$ with the subtle-obvious ratings of Christian et al.

There is evidence of a monotonic relationship between endorsement of an item and the social desirability value of the item (Edwards, 1953; Edwards & Walsh, 1963). Match and Wiggins (1974) found that the ability to estimate what is socially desirable is moderated by the subject's idiosyncratic perceptions, which are determined by the milieu from which the individual comes. Kimber (1947) concluded that these perceptions vary widely. Among those who have attributed systematic variance in the MMPI to social desirability are Edwards (1961), Edwards & Diers (1962), and Fordyce (1956) with Siller and Chipman (1963) finding that it accounted for a minimal amount of variance.

Whereas social desirability as a response set is related to item content, the set of acquiescence is dependent upon content and item characteristics. Acquiescence was imputed to be a major source of variance in the MMPI by Messick and Jackson (1961a) although Rorer (1965) found that it contributed little to the variance.

Bentler, Jackson, and Messick (1971) identified two components of acquiescence: true responding and item endorsement. True responding is the tendency to answer true, and item endorsement is the tendency to attribute characteristics to oneself regardless of the direction of the keying. To exemplify these definitions: The item "I work under a

great deal of tension" would be answered true by individuals exhibiting true responding and also by those exhibiting item endorsement, but the negation of this item; i.e., "I do not work under a great deal of tension" would be answered true only by those high on true responding; those high on item endorsement would answer false. These findings were corroborated by Morf and Jackson (1972) and disputed by Block (1971).

There has been much disagreement over the role that response sets play; however, it is generally acknowledged that they are maximally operant under conditions of ambiguity (Fiske, 1971, p. 215). Ambiguity has been defined relative to test instructions and test items. Fiske (1957b) theorized that response variability is reduced if the instructions minimize ambiguity. An investigation by Fink and Butcher (1972) failed to support Fiske's hypothesis although the variance was decreased for the cannot say response. Tracy and Fiske (1974) attempted to reduce response heterogeneity by giving subjects an explanation of the tests and allowing them the opportunity to ask questions prior to the testing. Their results led them to conclude that a subject's responses are determined by subject-item interaction rather than the experimenter's explanation.

Two dimensions of item ambiguity have been identified: interpretive, and response ambiguity (Broen, 1960; Goldberg,

1963). Interpretive ambiguity is the subject's perception of the item and has been measured subjectively by ratings (Harris & Baxter, 1965). Response ambiguity is the subject's perception of the item in relation to him- or herself and has been measured objectively by response changes within individuals over time (Fiske, 1957b). Baxter and Morris (1968) demonstrated that items high in interpretive ambiguity and low in response ambiguity have greater discriminating power. According to Goldberg (1963), there is more consistent response to items of extreme endorsement frequencies, which have been shown to be less ambiguous. However, when subjects were divided into sensitizers (i.e., those showing lowered thresholds for emotional stimuli) and repressors (i.e., those showing elevated thresholds for emotional stimuli), it was determined that sensitizers tend to make more frequent extreme responses in the presence of ambiguous stimuli (Norman, R. P., 1969).

Although Cronbach (1946, 1950) emphasized response sets as a source of cumulative error, hence a cause of interpretive invalidity, he did suggest that response tendencies could denote personality dynamics if these tendencies were consistent over time. Jackson and Messick (1958), in exploring Cronbach's postulate, used the term style to indicate that component of a response set which contributes to, rather than reduces, validity.

The conclusions drawn from these and other studies have ranged from the position that the major proportion of variance in the MMPI is due to the influence of response sets to one that maintains sets account for a negligible proportion. In most recent years, there has been comparatively little in the literature concerning response sets, a fact which has been attributed to the diversity of the findings with little hope of resolution (Taylor, Carithers, & Coyne, 1976).

Demand Characteristics

Some investigators have interpreted MMPI responding in terms of demand characteristics. Orne (1962) was the first to describe "the totality of cues which convey an experimental hypothesis to the subject [and] become significant determinants of the subjects' [sic] behavior. . . . [which he labeled] the demand characteristics of the experimental situation" (p. 779). The existence of this phenomenon became apparent to Orne when he attempted--and failed--to find some meaningless task which subjects would refuse to perform or continue to perform. His subjects were instructed to add long columns of figures and immediately tear the paper into small pieces; then they were told to add more columns of figures and again tear the paper. Contrary to Orne's expectations, the subjects repeated the process until instructed to stop.

Goldfried (1976) believed that demand characteristics

operate when self-report procedures, as well as situational and role-playing methods, are used. Harvey and Sipprelle (1976) attributed the results of their findings in an experimental setting to demand characteristics. In situ studies by MacLean, Tait, and Catterall (1953) and Drasgow and Bannette (1957), in which subjects were presumed to have the common motivational goal of creating a very favorable impression, demonstrate this concept also. Other investigators have concluded that implicit cues given to their subjects, in addition to the explicit instructions, enabled them to respond more accurately to the role demands of the experiment (Kroger, 1974; Kroger & Turnbull, 1970).

Studies in which the results are stated to be a function of demand characteristics infer that subjects are compliant in an experimental setting; however, there is evidence to the contrary. In testing subjects under three different instructional sets, Jackson and Messick (1969) found that some subjects confined all of their responses to the limitations imposed by only one of the sets. Fiske and Butler (1963) state that "the experimenter rarely has evidence to justify the inference that the criteria used by the subject in selecting his responses are solely those desired by the experimenter" (p. 254).

Impression Management

Goffman (cited in Braginsky, Braginsky, & Ring, 1969,

pp. 50-51) offers the explanation that, in some instances, individuals may be overtly compliant while subtlety engaging in impression management. Braginsky, Grosse, and Ring (1966) administered 30 MMPI items of neutral desirability value to two pairs of old timer-short timer hospitalized psychiatric groups. For one pair the test was labeled Mental Illness Test; for the other pair it was labeled Self-Insight Test. The instructions to both pairs were the same: marking an item true indicated congruence with the test title. As a control, a pair of old timer groups were administered the tests without the endorsement instructions. The investigators had hypothesized that the short timers, whose symptoms were not controlled but who wanted to leave the hospital, would manipulate their scores and present themselves as healthy. The reverse was hypothesized for the old timers. Both hypotheses were supported. Stricker (1969) also cites the function of impression management in role-playing studies.

Test-Wiseness

Test-wiseness is a factor which some writers have used to explain how individuals are able to produce profiles that do not reflect their own personality dynamics. Millman, Bishop, and Ebel (1965) define test-wiseness as

a subject's capacity to utilize the characteristics and the formats of the test and/or the test taking

situation to receive a high score. . . . [and] is logically independent of the examinee's knowledge of the subject matter for which the items are supposedly measures. (p. 707).

Numerous studies have been done to investigate test-wiseness relative to ability testing (cf. Millman et al.). Stricker (1969) studied the concept in relation to structured personality tests and concluded that test-wiseness measures were positively correlated with social desirability scales. Test-wiseness has been implicated in the ability to fake well by Kaufman (1950); however, Windle (1955) found changes in test scores (test-retest) to be unrelated to test-wiseness.

Faking Ability

In a discussion of faking, Kroger (1974) wrote, Success depends upon favorable motivation, . . . on an accurate conception of the role involved, . . . being skillful in acting as if one were someone else, . . . and on the presence of relevant cues in the test. (p. 131)

It is difficult, however, to determine the specific attributes that contribute to the ability to dissemble successfully because the results of much of the research are equivocal.

The ability to fake well was found to be unrelated to

IQ by Kimber (1947) and Noll (1951) and to be related to IQ by Cohn (1952). Verbal proficiency enabled subjects to fake well (Levitt & Rice, 1968) and did not appear to be helpful (Burkhart, Gynther, & Christian, 1978). Sex was not found to be a factor in successful dissembling (Levitt & Rice, 1968; Parsons, Yourshaw, & Borstelmann, 1968); other research has indicated it was a factor (Kimber, 1947; Merbaum, 1972). Test anxiety had a positive effect on the ability of some subjects to fake (Baldry & Sarason, 1968) and a negative effect on others (Hoffman, 1968). According to Welsh and Dahlstrom (cited in Lawton, 1963), subjects who are psychologically sophisticated are better able to simulate; however, Burkhart, Gynther, and Christian (1978) did not find this to be true.

Two attributes contributing to successful dissembling are not disputed: psychological adjustment and motivation. Anthony (1976) believes that "psychopathology may interfere with successful simulation of healthy roles" (p. 39). When Grayson and Olinger (1957) asked psychiatric patients, all of whom had deviant T-scores, to simulate, only 11% were able to produce normal profiles although 73% showed some improvement; the remaining subjects actually showed greater pathology. Other studies also indicated that abnormal subjects are less able to dissimulate successfully than normal subjects (Lanyon, 1967; Lawton & Kleban, 1965; Milgram,

1960).

Investigating the effects of motivation upon dissimulation, Heron (1956) tested two groups of job applicants. One group was tested during the hiring process, and the second group was tested after they had been told they were hired. The profiles of the subjects in the first group indicated better adjustment leading Heron to conclude that they were more motivated to make a good impression. Similar studies with normal (Mayo & Guttman, 1959) and abnormal (Wilcox & Krasnoff, 1967) subjects led to the same conclusion.

Effects of Faking on Variance

Fiske (1957a) believes that "low variability [in test scores] indicates a well-structured orientation toward the given stimuli" (p. 464). That well structured orientation is often attributed to the effects of the instructional set, which has been shown in the many studies on response sets. Parsons et al. (1968) found that their results of successful ideal-self responding were due to a social desirability response set induced by the instructions. Burkhart, Gynther, and Christian (1978) also believed their results were due to the effects of the instructional set. Sheldon (1959) found that his subjects were more successful when hints on responding were given.

Studies designed to elicit fake good or ideal-self responding in normal subjects demonstrate that the variance is

significantly less than it is under honest conditions (Binder, Mayman, & Doehrman, 1974; Exner, McDowell, Pabst, Stackman, & Kirk, 1963; Rapaport, 1958; Voas, 1958; Wales & Seeman, 1968, 1972). However, in studies in which subjects, normal or abnormal, were asked to fake bad, the variance is greater than under conditions of honest responding (Anthony, 1971; Branca & Podolnick, 1961; Gough, 1947; Wales & Seeman, 1969). Gough (1950) believed this fact reflected the inability of subjects to approximate profiles of those whose personality dynamics they do not understand.

Motivation

Cattell, Eber, and Tatsuoka (1970, chap. 6) distinguish among four types of deception, each arising from a different source: sabotage, self-illusion, motivational distortion, and cooperative faking. It is the latter two that are of concern in the studies described above. Motivational distortion, whether conscious or unconscious, arises from an internal stimulus: it is self-motivation to deceive, whereas cooperative faking is the conscious deception that results from subjects aligning their responses with an external criterion. The two differ in origin; therefore, it can not be assumed that they produce equivalent results.

Hunt (1948) acknowledged that the results of his investigation of deliberate deception could not be generalized to populations in which the deception is surreptitious. Dahl-

strom and Welsh (1960, p. 137) did not consider results obtained by instructing subjects to fake bad to be equivalent to those obtained when subjects attempted to malingering independent of instructions. Cooke and Robey (1971) concluded that "the difference in the nature of the motivation between an external instructional motivation and an internally driven need to dissimulate may lead to differing effects on the MMPI" (p. 359).

Cognitive vs. Affective Responding

It has long been recognized that words have specific as well as general meaning for people. In Osgood's (1952, 1962) work on the semantic differential, he distinguished between the cognitive dictionary definition of words and the affective meaning that they have for individuals. The numerous investigations of perceptual defense in the 1940's and 1950's demonstrated that subjects react differently toward emotionally toned stimuli than they do toward neutral stimuli and that emotionally toned stimuli are person-specific (Blum, 1955; Bruner & Postman, 1947; Erickson & Browne, 1956; Mathews & Wertheimer, 1958). I believe that when subjects respond to personality inventory items, they do so affectively according to the affective meaning that the items have for them.

According to Singer and Young (1941), subjects make cognitive evaluations and judgments in the absence of affective

arousal. Many MMPI studies, including those investigating the properties of the zero-X and subtle-obvious dimensions, typically are conducted by testing different groups of subjects under various instructional sets; i.e., by giving general instructions to fake good or bad, or by requesting subjects to assume a certain role. I maintain that overtly enlisting the subject's cooperation to fake diminishes or extinguishes affective arousal and causes the subject to respond according to the cognitive meaning of the item as defined by the particular instructional set. In effect, it causes test taking to become a cognitive process--a test of the subject's ability to dissemble--rather than an affective process. Results thus obtained do not reflect the heterogeneity associated with the subject's own personality dynamics, which is reflected when the subject responds honestly or when he or she is self-motivated to deceive.

A purpose of this study is to demonstrate the decrease in variance that results when subjects are tested under increasingly explicit instructions to fake. In the experimental setting, it is difficult to induce deception that is self-motivated because of ethical considerations; however, it is theorized that implicit instructions to fake can have the same effect, and that the results will indicate subjective and not objective responding.

Method

Subjects

Two hundred and thirty-nine volunteers from undergraduate psychology classes at Texas Woman's University and Henderson County Junior College were randomly assigned to four groups. Informed consent was obtained from all subjects in accordance with APA principles. The subjects were told that the purpose of the research was to examine group responses to MMPI items. The subjects received extra credit from their instructors for participation in the study.

It was necessary to discard 12 protocols because of randomness in responding or an insufficient number of items completed (i.e., fewer than 536). Of the remaining 227 subjects, 47 were males and 180 were females with a mean age of 22.0 and 22.5 years, respectively.

Procedure

The subjects were administered the 566 items on the Group Form Booklet of the MMPI. Each subject was given the test booklet with the instructions stapled to the cover. The subjects were told to read the instructions carefully and to refer to them during the session to ensure that they were following them correctly. Each group was tested under a different instructional set. Appendix A contains the com-

plete instructions for each group.

Group 1 (honest responding) was tested under the standard instructions, which are that the responses are to be the subject's own opinion of him- or herself. These subjects were instructed to fill in their sex and age but not their names. It was expected that this anonymity would increase the subject's tendency to respond honestly.

Group 2 (self-motivated deception, implicit) was tested under the standard instructions also; however, they were instructed to sign their names. The assumption was made that the loss of anonymity would motivate subjects to present themselves more favorably than they would if their identity were known.

Group 3 (self-motivated deception, explicit) was given a brief description of a well-adjusted individual with respect to MMPI parameters and was told that the profiles would be compared with that of a well-adjusted person. The subjects in this group were also instructed to sign their names. It was expected that these instructions would motivate the subjects to fake good.

Group 4 (cooperative faking) was tested under explicit instructions to fake good; i.e., the subjects were instructed to respond, not as they believed themselves to be, but as they would if it were imperative for them to make a very good impression. The subjects in this group were also instructed

to sign their names. It was expected that these instructions would cause the subjects to align their responses with an external criterion; e.g., social desirability.

All of the protocols were scored with the standard keys to obtain T-scores with K-corrections added for the three validity scales and nine of the clinical scales. (Mf/Masculine-feminine was not included because the protocols were not differentiated by sex in the analysis.) The protocols were also scored with special zero-X keys to obtain zero and X item scores for three subtlety categories using the S-O ratings of Christian et al. (1978). The responses for somewhat obvious and very obvious items were not scored because there is only one zero item in the former category, and there are none in the latter. Each zero-X item was scored only once irrespective of the number of times it is tabulated for the standard scales. The raw scores for zero-X items were converted to proportions for use in the analyses due to the unequal base rate of the items. There are 32, 39, 12, 23, 61, and 132 items in the zero very subtle, zero somewhat subtle, zero neutral, X very subtle, X somewhat subtle, and X neutral categories, respectively, for males. The number of items is the same for females with the exception of the latter two categories which have 64 and 130 items, respectively.

Results

The means and standard deviations for the zero and X item scores according to degree of subtlety for each instructional set group are presented in Table 2. A discriminant function analysis was done with Instructional Set as the criterion variable and the six zero-X measures (ZVS, ZSS, ZN, XVS, XSS, XN) as the predictor variables. Contrary to expectations, none of the preliminary one-way analyses of variance in this analysis yielded a significant F ; the largest obtained value was for XVS, $F(3, 223) = 1.216$, $p < .31$. These results indicate that the instructional set groups could not be differentiated on the basis of frequency of endorsement of zero-X items.

In order to group the protocols according to how "good" the subjects presented themselves on the inventory, three MMPI indices were considered: the F and K scales and Gough's F-K dissimulation index. As discussed above, F and K are unipolar scales which measure different dimensions of denial or exaggeration of psychopathology. Each of these scales was rejected as the grouping factor because they can produce contradictory results; e.g., a subject may have a high score on F , indicating admission of psychopathology and, at the same time, have a high score on K indicating subtle

Table 2

Means and Standard Deviations for Proportions
of Zero-X Items Endorsed Within Each Subtlety
Category for Instructional Set Groups

		Honest	Implicit	Explicit	Fake Good
		<u>n</u> = 58	<u>n</u> = 57	<u>n</u> = 56	<u>n</u> = 56
ZVS	\bar{X}	.6579	.6897	.6981	.6870
	SD	.1437	.1184	.1353	.1451
ZSS	\bar{X}	.5694	.5551	.5719	.6062
	SD	.1644	.1592	.1570	.1787
ZN	\bar{X}	.5014	.5146	.5313	.5015
	SD	.1282	.1347	.1381	.1060
XVS	\bar{X}	.4730	.4798	.4829	.5186
	SD	.1491	.1447	.1096	.1497
XSS	\bar{X}	.3602	.3471	.3573	.3680
	SD	.0648	.0780	.0728	.0758
XN	\bar{X}	.3393	.3117	.3086	.3070
	SD	.1168	.0936	.0854	.1237

denial or defensiveness. The F-K index corrects for this contradiction and was therefore used as the grouping factor.

Numerous F-K cutting scores have been proposed in the literature to differentiate malingered and dissimulated profiles from those of normals (cf. Gough, 1950; Hunt, 1973). Four cutting scores which are frequently cited are ± 12 , ± 11 , ± 10 , and ± 9 . The cutting scores used to group the protocols for subsequent analyses were ± 12 because these scores result in a more equal n across the groups. Each protocol was assigned to one of the following four groups according to F-K score: (1) at or above the upper cutting score, (2) $+1$ to the upper cutting score, (3) -1 to the lower cutting score, and (4) at or below the lower cutting score. There were four protocols with zero F-K scores, and these were omitted from the analyses. The resulting four groups were named, in order, Fake Bad, Look Bad, Look Good, and Fake Good. There were 11, 29, 99, and 84 subjects in the four respective groups.

The means and standard deviations for the six zero-X categories for the four groups are shown in Table 3. The contrast between zero and X items of equal degrees of subtlety within each group are illustrated in Figure 1. Twelve correlated samples t tests were done to test for differences between these means. Each of the differences was tested at $\alpha = .038$ to maintain a Type I risk of .05. No significant

Table 3

Means and Standard Deviations for Proportions of Zero-X
Items Endorsed Within Each Subtlety Category for
Groups Formed According to F-K Scores

		Fake Bad ^a	Look Bad ^b	Look Good ^c	Fake Good ^d
		<u>n</u> = 11	<u>n</u> = 29	<u>n</u> = 99	<u>n</u> = 84
ZVS	\bar{X}	.4887	.5388	.6727	.7768
	SD	.0865	.0867	.1154	.0945
ZSS	\bar{X}	.3939	.3873	.5343	.7186
	SD	.1602	.0887	.1113	.1158
ZN	\bar{X}	.4546	.4310	.5051	.5576
	SD	.1413	.0973	.1248	.1227
XVS	\bar{X}	.4980	.4828	.4901	.4912
	SD	.1667	.1539	.1355	.1355
XSS	\bar{X}	.4425	.3780	.3611	.3386
	SD	.0973	.0672	.0732	.0617
XN	\bar{X}	.5046	.4379	.3293	.2326
	SD	.0742	.0794	.0714	.0671

^aF-K > 11.

^b12 > F-K > 0.

^c0 > F-K > -12.

^dF-K < -11.

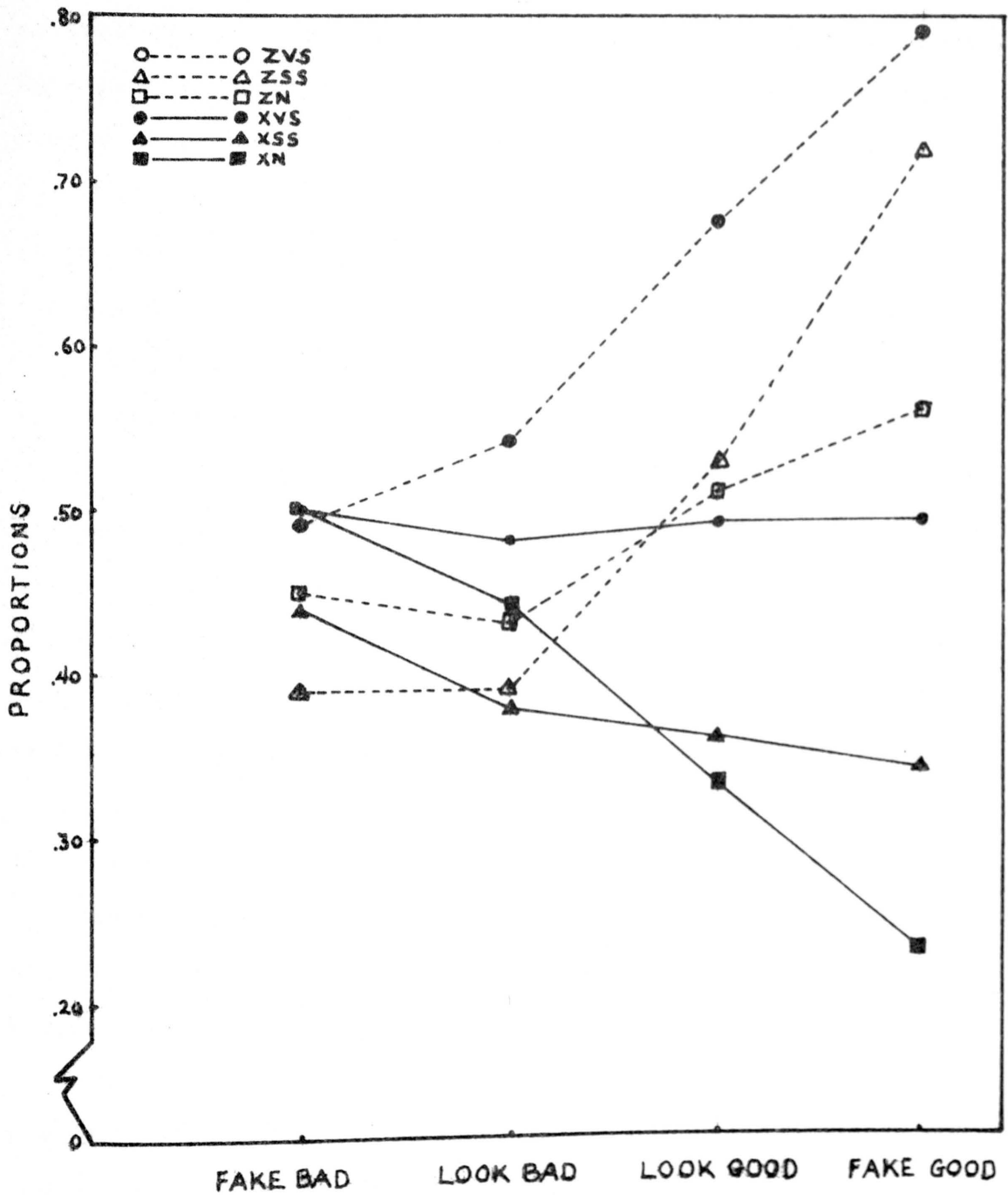


Figure 1. Mean scores on ZVS, ZSS, ZN, XVS, XSS, and XN items for Fake Bad, Look Bad, Look Good, and Fake Good groups.

differences were found for the Fake Bad, Look Bad, or Look Good groups; however, the zero scores were significantly higher than the X scores for all three subtlety categories for the Fake Good group, $p < .05$. These results are presented in Table 4. Further analysis was necessary to help clarify this finding.

Multiple regression was used to determine the degree of relationship between the six categories of zero-X items and three validity indices: the K scale, as a measure of defensiveness; the F scale, as a measure of psychopathology; and the F-K index, as a bipolar measure with psychopathology minus defensiveness for positive scores and defensiveness minus psychopathology for negative scores. The analysis of variance for each of the three regression equations shows that, as a group, the six types of zero-X items account for a significant proportion of variability in the three criterion measures, $p < .001$. These analyses are summarized in Table 5. The standard regression method with stepwise inclusion of the variables was used.

Partial F tests were done for each of the analyses to test for significant changes in R^2 . The level of significance was set at .04 to hold the Type I risk at .05. Approximately 52% of the variance on the F scale is due to these more subtle items, and over 46% of this variance is due to X neutral items (see Table 6). X very subtle items, and zero

Table 4
Mean Differences and \underline{t} Values on Zero-X Items
for Fake Bad, Look Bad, Look Good,
and Fake Good Profiles

		VS Z-X		SS Z-X		N Z-X	
	<u>df</u>	<u>D</u>	<u>t</u>	<u>D</u>	<u>t</u>	<u>D</u>	<u>t</u>
Fake Bad	10	-.0093	-.06	-.0486	-.58	-.0500	-.31
Look Bad	28	.0560	.32	.0093	.10	-.0069	-.07
Look Good	98	.1826	1.10	.1732	1.36	.1758	1.13
Fake Good	83	.2856	2.25*	.3800	3.07*	.3250	2.27*

* $p < .05$.

Table 5

Analysis of Variance Summaries for Three Regression
Equations with the K Scale, the F Scale,
and the F-K Index as Criterion Measures

	SS	<u>df</u>	MS	<u>F</u>
<u>K</u> Scale				
Regression	13,384.87	6	2,230.81	121.7631*
Residual	3,957.32	216	18.32	
<u>F</u> Scale				
Regression	19,607.81	6	3,267.97	38.8068*
Residual	18,189.64	216	84.21	
F-K Index				
Regression	13,778.32	6	2,296.39	105.5194*
Residual	4,700.74	216	21.76	

* $p < .001$.

Table 6

Significance Tests of Regression Coefficients
with the F Scale as the Criterion

Variable	Multiple <u>R</u>	<u>R</u> ²	<u>R</u> ² Change	<u>F</u> ^a
XN	.68276	.46616	.46616	209.0404**
ZVS	.70593	.49834	.03218	14.4305**
XSS	.71816	.51575	.01740	7.8027*
XVS	.71984	.51817	.00242	1.0852
ZSS	.72014	.51861	.00044	.1973
ZN	.72025	.51876	.00015	.0679

^adf = 1, 216.

*p < .01.

**p < .001.

somewhat subtle and neutral items do not cause a significant increase in the multiple correlation coefficient. Results concerning the K scale (see Table 7) show that 77% of the variance is accounted for by these six types of items with 72% attributed to zero somewhat subtle items. Very subtle items, both zero and X, caused no significant increase in the multiple correlation coefficient. Three categories of items yielded significant F tests for the F-K index: X neutral items, and zero somewhat subtle and very subtle items, in order of magnitude (see Table 8). Less than .5% of the total variance accounted for is due to the addition of the remaining three item types.

The specific contribution of the six types of zero-X items to psychopathology and defensiveness can be understood better by an examination of the Beta weights, which are presented in Table 9. These results show that X neutral items provide the single best index of psychopathology among these six types of items; whereas, zero neutral items only contribute to defensiveness. The findings for X somewhat subtle items appear to be contradictory; these items make significant contributions to both psychopathology and defensiveness. Zero somewhat subtle items are associated with defensiveness only. X very subtle items have a negligible effect with respect to both psychopathology and defensiveness. Endorsement of zero very subtle items is concordant with negative

Table 7

Significance Tests of Regression Coefficients
with the K Scale as the Criterion

Variable	Multiple <u>R</u>	<u>R</u> ²	<u>R</u> ² Change	<u>F</u> ^a
ZSS	.84872	.72032	.72032	686.0190**
ZN	.86379	.74613	.02580	24.5714**
XN	.86987	.75667	.01054	10.0381*
XSS	.87629	.76789	.01122	10.6857**
XVS	.87819	.77122	.00333	3.1714
ZVS	.87853	.77181	.00059	.5619

^adf = 1, 216.

*p < .002.

**p < .001.

Table 8

Significance Tests of Regression Coefficients
with the F-K Index as the Criterion

Variable	Multiple <u>R</u>	<u>R</u> ²	<u>R</u> ² Change	<u>F</u> ^a
XN	.80659	.65059	.65059	556.0598*
ZSS	.84826	.71955	.06896	58.9402*
ZVS	.85918	.73818	.01863	15.9231*
ZN	.86160	.74236	.00417	3.5641
XVS	.86239	.74372	.00136	1.1624
XSS	.86349	.74562	.00190	1.6239

^adf = 1, 216.

*p < .001.

Table 9

Beta Weights and Multiple Correlation Coefficients
for Three Criterion Variables

Scale	Multiple <u>R</u>	ZVS	ZSS	ZN	XVS	XSS	XN
<u>F</u>	.720	-.259	.030	.013	-.053	.156	.469
<u>K</u>	.879	-.039	.643	.181	.069	.108	-.257
F-K	.863	-.180	-.262	-.070	-.150	.054	.470

F-K scores and low scores on F.

The means and standard deviations for the 12 MMPI scales on the Fake Bad, Look Bad, Look Good, and Fake Good profiles are presented in Table 10. On the clinical scales, both of these measures decrease with increasing attempts to look good with the exception of the standard deviation for Si. On the validity scales, the mean on K increases as the subjects present themselves more favorably and the mean on F-K decreases, which would be expected. On L, both the mean and standard deviation decrease as F-K scores approach zero, then they increase as F-K becomes negative.

An overall Levene's test on the T-score variances for each scale indicates extreme non-homogeneity for all scales with the exception of Ma (see Table 11). Because 12 comparisons were made, the level of significance was set at .038 to maintain a Type I risk of .05.

To test the prediction that variability decreases as subjects attempt to make a good impression, the equality of variances for the T-scores was tested, using Levene's test with $\alpha = .025$ for the 33 comparisons, comparing Fake Bad profiles with those of Look Bad, Look Good, and Fake Good, in order. The variances for these pairwise comparisons are presented in Table 11. The number of significant differences found increases with increased attempts to look good. Where there are significant differences, the smaller vari-

Table 10

Means and Standard Deviations of MMPI Scales
on Fake Bad, Look Bad, Look Good,
and Fake Good Profiles

Scale		Fake Bad	Look Bad	Look Good	Fake Good
		<u>n</u> = 11	<u>n</u> = 29	<u>n</u> = 99	<u>n</u> = 84
<u>L</u>	\bar{X}	51.73	44.93	46.90	52.46
	SD	13.47	5.50	6.93	9.29
<u>F</u>	\bar{X}	98.91	70.76	56.32	49.36
	SD	14.77	8.79	5.01	3.98
<u>K</u>	\bar{X}	44.73	41.62	50.19	61.13
	SD	12.38	4.47	4.80	5.28
<u>Hs</u>	\bar{X}	67.27	53.69	50.26	50.11
	SD	19.24	10.53	8.66	6.43
<u>D</u>	\bar{X}	73.64	56.21	49.44	46.75
	SD	19.65	10.26	9.37	6.93
<u>Hy</u>	\bar{X}	62.18	53.79	52.52	55.44
	SD	18.31	10.36	8.26	7.94
<u>Pd</u>	\bar{X}	73.64	65.10	58.21	58.13
	SD	13.46	12.06	9.13	8.36
<u>Pa</u>	\bar{X}	78.18	64.45	55.62	53.12
	SD	16.91	9.01	8.85	7.46
<u>Pt</u>	\bar{X}	79.09	64.69	56.27	53.39
	SD	13.47	10.76	9.74	6.67

<u>Sc</u>	\bar{X}	94.91	75.48	59.08	55.41
	SD	16.81	16.07	10.38	6.94
<u>Si</u>	\bar{X}	65.18	59.79	53.35	46.18
	SD	7.18	6.48	9.28	7.79
F-K ^a	\bar{X}	16.73	4.24	-6.85	-15.79
	SD	5.90	2.91	2.99	3.01

^aRaw scores.

Table 11

Levene's Test for Overall Equality of Variances of MMPI
Scales and Pairwise Variance Comparisons Between
Fake Bad Profiles and Look Bad, Look Good,
and Fake Good Profiles, Respectively

Scale	Variances				
	Overall ^a	F-B	L-B ^b	L-G ^c	F-G ^d
<u>L</u>	74.49*	181.44	30.25*	48.02*	86.30
<u>F</u>	170.20***	218.15	77.26**	25.10***	15.84***
<u>K</u>	78.85***	153.26	19.98***	23.04***	27.88***
<u>Hs</u>	93.61***	370.18	110.88*	75.00***	41.34***
<u>D</u>	123.99***	386.12	105.27	87.80***	48.02***
<u>Hy</u>	87.35***	335.26	107.33	68.23***	63.04***
<u>Pd</u>	104.73*	181.17	145.44	83.36	69.89
<u>Pa</u>	114.49**	285.95	81.18	78.32*	55.65**
<u>Pt</u>	118.14***	181.44	115.78	94.87	44.49**
<u>Sc</u>	208.37***	282.58	258.24	107.74	48.16***
<u>Ma</u>	136.68	147.06	153.02	138.70	79.98
<u>Si</u>	98.53*	51.55	41.99	86.12	60.64

^adf = 3, 219.

^bdf = 1, 38.

^cdf = 1, 108.

^ddf = 1, 93.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

ance is associated with the group trying to make the better impression. (These results do not reflect the significant differences found for Pd and Si in pairwise comparisons not used in this study.)

Discussion

Instructional Sets

Because the groups derived on the basis of F-K index scores could be differentiated with respect to frequency of endorsement of zero-X items and those formed by instructional sets could not be, it is concluded that the instructional sets failed. It is acknowledged that the two sets of groups would not be equivalent had the instructions differentiated the groups; e.g., all of the subjects in instructional set Group 1 (honest responding) would not have had F-K scores greater than 11. The failure is hypothesized to be due to two confounding factors: deficits inherent in the instructions and subject-selection bias.

The standard instructions used for Group 1 and the explicit instructions to fake good used for Group 4 (cooperative faking) have been used successfully in many previous studies assessing change scores on zero-X and subtle-obvious items (Burkhart, Gynther, & Christian, 1978; Wales & Seeman, 1972). It is therefore assumed that, for the majority of subjects, these instructions were understandable and not subject to misinterpretation. However, it is a real possibility that the instructions for Groups 2 and 3, self-motivated deception--implicit and explicit, respectively, were

faulty. With respect to instructions for the former group, it may be that loss of anonymity, in past years considered an inducement for subjects to present themselves more favorably, is no longer the motivating factor it once was. A possible explanation for the failure of the instructional set for Group 3 is that subjects may have been unable to find parallels between specific items and characteristics of a well-adjusted person.

Although the subjects were told that participation in the study was voluntary, all received extra credit and nearly 70% were tested during a scheduled class period. Perhaps these inducements to participate placed constraints upon some subjects to which they reacted by being inattentive to the instructions or by responding in opposition to the instructions. It became evident from the behavior of some subjects during the session and as the protocols were turned in that the instructions were read carelessly or not at all; e.g., questions were asked that would not have been had the instructions been read; some subjects had to be reminded to fill in the demographic data; some who had been instructed to sign their names failed to do so, and the reverse was also true. There were obvious attempts at sabotage. Approximately 16% of the subjects had one or more T-scores in a range indicative of malingering and/or produced profiles that would seldom be found outside of a clinical setting.

In Groups 3 and 4, 9.7% of the profiles were one of these types.

Unfortunately, despite the examiner's intention and pretest explanation, some subjects were under the impression that they would receive interpretations of their individual results. This was evident upon completion of the testing when numerous inquiries were made about how and when the results would be available.

Any of these factors affecting the subject's motivation could cause responses to be biased. Although some subjects undoubtedly were willing volunteers who cooperated with the instructions to the best of their ability, each of the instructional set groups is considered to have resulted in valid and invalid profiles. If I were to replicate this study, I would try to use subjects who appear motivated to participate in research, test in smaller groups by instructional set and review the instructions verbally prior to the testing.

A purpose of this research was to demonstrate that variability decreases as subjects are tested under increasingly explicit instructions to fake. However, it is not possible to assess these results in terms of conscious and unconscious deception. All of the profiles in the Fake Bad group and many of the profiles in the Look Bad group were probably malingered. Whereas many of the profiles in the Fake Good

group were the result of overt attempts to "fake good," others appear to be due to honest responding.

The results are consistent with theories of variability in personality testing (Fiske, 1957b). As increased attempts were made to look good, variability decreased as parameters of social desirability were approximated, which has been shown in previous studies (Rapaport, 1958; Voas, 1958; Zimmerman & Gloye, 1959). It should be noted that the comparisons in these studies were between "real self" and "ideal self." On the clinical scales, only Pd, Ma, and Si exhibited homogeneity of variance. The latter two scales contain fewer obvious items than the other clinical scales; i.e., 23% and 17%, respectively. Therefore, the choice between acceptable and unacceptable responses is less clear and more guesses have to be made within all groups. This explanation does not extend to Pd, however, where 38% of the items are in the obvious ranges. The nonsignificant difference for L between the Fake Bad and the Fake Good groups is attributed to an increase in the denial of common flaws on the part of subjects trying to make a good impression. The results for L, Pd, and Ma (Si was not reported) are consistent with those given by Rapaport and Voas.

For groups of normals, used in this study, the approximation of deviant behaviors is a less common orientation than the recognition of socially desirable behaviors. This

is believed to account for the greatly increased variances for the Fake Bad group, which probably caused more of a bias in the direction of significance than would have existed had this group been comprised of honest responders.

Zero-X Items

Partial support was given to the hypothesis that there is no significant difference between zero and X items of equal degrees of subtlety. The means were equal until the F-K index approached $-.12$, at which point there were significant differences for all three levels of subtlety. The interpretation of the findings is predicated on the assumption that two main components of the MMPI are psychopathology and defensiveness, and that the latter may be indicative of pathology or nonpathological adjustment.

Among the six types of items analyzed, X neutral items seem to have the greatest discriminating power with respect to psychopathology and ability to compensate. This is consistent with findings that items whose content relevance is more apparent have greater discriminating power (Brozek & Erickson, 1948; Duff, 1965; Gloye & Zimmerman, 1967; McCall, 1968).

X very subtle items, on the other hand, have the least ability to discriminate. Endorsement of these items does not affect indices of psychopathology or defensiveness. An examination of these items reveals their relationship to

either index to be obscure. Most of these items pertain to vocational and social interests; e.g., "I like mechanics magazines," (scored false for both males and females). It may be that the inclusion of these items in the inventory was the result of item sampling or that changing mores over the years have diminished their validity. Whatever the case may be, the current results indicate that X very subtle items are irrelevant to criteria they are intended to measure, and that their contribution to the validity of the MMPI is extremely questionable.

Results concerning X somewhat subtle items appear to be diametric; i.e., endorsement of these items increases indices of psychopathology and defensiveness. These results are concordant with the original validation studies which found that subtle items can discriminate psychopathology and discrepant with studies that concluded that endorsement of subtle items may be indicative of nonpathological responding (Burkhart, Christian, & Gynther, 1978; Burkhart, Gynther, & Christian, 1978; Rosen, 1956).

Endorsement of zero very subtle items seems to be consistent with psychological integration; evidence of psychopathology is decreased with no significant increase in defensiveness. These results, together with those for X very subtle items, strongly suggest that there is a minimum level of subtlety for MMPI items below which there is little or no

discriminating power for diagnostic criteria. This finding is in support of the conclusion drawn by Duff (1965) in his study of item subtlety and validity.

Zero somewhat subtle and zero neutral items have their greatest impact on defensiveness with a negligible effect on psychopathology. The dynamic meaning of defensiveness on the MMPI, as measured by the K scale, is confounded by the direction of scoring of the items. Of the 30 K items, 29 are scored false. For the six categories used in this study, 51.6% of the X items are scored in the negative direction as opposed to 86.8% of the zero items. Zero-order correlations between X items and K are negative; whereas, the reverse is true for zero items. Endorsement of zero items may reflect a response tendency to answer false. Fricke (1957) hypothesized a similar relationship for subtle items, many of which are scored in the negative direction. The order of magnitude for zero-order correlations between the items and K, however, is not related to the number of items in a given category and the number of items appearing on K. It may be that items with negative Beta weights with respect to the F-K index (see Table 9) are more unbiased indices of defensiveness.

The relationship of F-K to defensiveness can be better understood by examining changes in F-K scores and F scores with respect to double-scored items; i.e., those items

scored for psychopathology whether answered true or false. There are 38 of these items among the zero items, and they are listed in Table 12. (Items scored opposite for males and females are not included.) It will be noted that there is considerable overlap between X neutral items and zero items. If a subject who endorsed a given number of these double-scored items in the X direction changed the scoring to the zero direction so that somewhat subtle and/or neutral items were endorsed, psychopathology would be decreased because of the decreased number of X items endorsed. The F-K index score would reflect a move away from psychopathology toward defensiveness because of the decrease in the number of X items endorsed and the increase in the number of zero items endorsed.

The question is, does defensiveness on the MMPI represent pathology or nonpathological adjustment? Increased endorsement of zero items under fake good conditions was found in normal populations (Wales & Seeman, 1968; 1972) and in abnormal populations (Anthony, 1971; Vesprani & Seeman, 1974; Wales & Seeman, 1969). The current results indicate that endorsement of zero items in a normal population denotes psychological integration. It is hypothesized that endorsement of zero items in an abnormal population indicates recognition of normal behaviors not necessarily found in the individual but with which he or she tries, consciously or uncon-

Table 12

Zero-X and Subtle-Obvious Classifications
for Double-Scored MMPI Items

Item	Classification							
	ZVS	ZSS	ZN	XVS	XSS	XN	ZO	XO
15 ^a						F	T	
25	F			T				
33 ^b		F						T
64		F				T		
82		F				T		
96	T					F		
109		F				T		
111		T				F		
117		F				T		
124		F				T		
126		F		T				
130	T					F		
134	T					F		
138			F			T		
142		F				T		
143		F			T			
147		F				T		
171		F			T			
172		F				T		
180		F				T		

187	T			F	
193	T			F	
201		F		T	
217	F				T
233	F				T
241		F		T	
263	F			T	
267	F			T	
268		T		F	
271	F				T
279	F			T	
292	F			T	
314 ^a		T		F	
316		F		T	
323 ^b	F			T	
359	F			T	
383		F	T		
398	F				T

^aduplicate item.

^bduplicate item.

sciously, to identify. This rationale implies that endorsement of items considered defensive can lead to spurious results on the MMPI for normal subjects.

In summary, endorsement of X neutral items seems to indicate admission of psychopathology with decreased ability to compensate. Very subtle items on both zero and X dimensions are questionable in their ability to discriminate psychopathology, and it therefore appears plausible that they discriminated unknown criteria in the original validation studies. Contrary to studies done since the publication of the MMPI, X somewhat subtle items seem to be identified with pathology. Zero somewhat subtle and zero neutral items do not appear to be valid indices of psychopathology.

One implication of this study is that investigations assessing change scores on subtle-obvious items without regard for their zero-X classification can lead to distorted results.

Appendix A

Instructions to Group 1 (honest responding):

This inventory consists of numbered statements. Read each statement and decide whether it is true as applied to you or false as applied to you.

You are to mark your answers on the answer sheet which is in the booklet. If a statement does not apply to you or if it is something that you don't know about, make no mark on the answer sheet.

Remember to give YOUR OWN opinion of yourself. Do not leave any blank spaces if you can avoid it.

Fill in only your age and sex on the answer sheet. DO NOT SIGN YOUR NAME or any of the other demographic data.

Instructions to Group 2 (self-motivated deception, implicit):

This inventory consists of numbered statements. Read each statement and decide whether it is true as applied to you or false as applied to you.

You are to mark your answers on the answer sheet which is in the booklet. If a statement does not apply to you or if it is something that you don't know about, make no mark on the answer sheet.

Remember to give YOUR OWN opinion of yourself. Do not leave any blank spaces if you can avoid it.

Fill in your age, sex, and name on the answer sheet.

Please print legibly. REMEMBER TO FILL IN YOUR NAME.

Instructions to Group 3 (self-motivated deception, explicit):

This inventory consists of numbered statements. Read each statement and decide whether it is true as applied to you or false as applied to you.

You are to mark your answers on the answer sheet which is in the booklet. If a statement does not apply to you or if it is something you don't know about, make no mark on the answer sheet. Do not leave any blank spaces if you can avoid it. Fill in your age, sex, and name on the answer sheet.

Your scores will be compared with those of well-adjusted individuals whose MMPI profiles indicate the following:

1. self-confident enough to admit to minor faults and defects
2. few bodily complaints and little concern about health
3. cheerful, enthusiastic, optimistic, active, and outgoing
4. a reasonable level of conformity to social regulations
5. an average female (male) interest pattern
6. adequate regard and interest for others
7. no undue worry or self-doubt
8. normal energy and activity level
9. a capacity to maintain rewarding relationships with others

REMEMBER TO FILL IN YOUR NAME ON THE ANSWER SHEET.

Instructions to Group 4 (cooperative faking):

This inventory consists of numbered statements to be answered true or false. Read each statement and decide if it would be true as applied to you or false as applied to you IF IT WERE IMPERATIVE FOR YOU TO MAKE A VERY GOOD IMPRESSION. Do not give your own opinion of yourself--respond as you would if you had to make a very good impression. Fill in your age, sex, and name on the answer sheet.

References

- Anthony, N. Comparison of clients' standard, exaggerated, and matching MMPI profiles. Journal of Consulting and Clinical Psychology, 1971, 36, 100-103.
- Anthony, N. Malinger as role taking. Journal of Clinical Psychology, 1976, 32, 32-41.
- Baldry, A. I., & Sarason, I. G. Test anxiety, preliminary instructions, and responses to personality inventories. Journal of Clinical Psychology, 1968, 24, 65.
- Baxter, J. C., & Morris, K. L. Item ambiguity and item discrimination in the MMPI. Journal of Consulting and Clinical Psychology, 1968, 32, 309-313.
- Bentler, P. M., Jackson, D. N., & Messick, S. Identification of content and style: A two-dimensional interpretation of acquiescence. Psychological Bulletin, 1971, 76, 186-204.
- Berg, I. A. Research notes from here and there Deviant responses and deviant people: The formulation of the deviation hypothesis. Journal of Counseling Psychology, 1957, 4, 154-161.
- Binder, J., Mayman, M., & Doehrman, S. Self-ideal-self discrepancy as a defensive style. Comprehensive Psychiatry, 1974, 15, 335-343.

- Block, J. On further conjectures regarding acquiescence. Psychological Bulletin, 1971, 76, 205-210.
- Blum, G. S. Perceptual defense revisited. Journal of Abnormal and Social Psychology, 1955, 51, 24-29.
- Braginsky, B. M., Braginsky, D. D., & Ring, K. Methods of madness: The mental hospital as a last resort. New York: Holt, Rinehart and Winston, Inc., 1969.
- Braginsky, B. M., Grosse, M., & Ring, K. Controlling outcomes through impression-management: An experimental study of the manipulative tactics of mental patients. Journal of Consulting Psychology, 1966, 30, 295-300.
- Branca, A. A., & Podolnick, E. E. Normal, hypnotically induced, and feigned anxiety as reflected in and detected by the MMPI. Journal of Consulting Psychology, 1961, 25, 165-170.
- Broen, W. E. Ambiguity and discriminating power in personality inventories. Journal of Consulting Psychology, 1960, 24, 174-179.
- Brozek, J., & Erickson, N. K. Item analysis of the psychoneurotic scales on the Minnesota Multiphasic Personality Inventory in experimental starvation. Journal of Consulting Psychology, 1948, 12, 403-411.
- Bruner, J. S., & Postman, L. Tension and tension release as organizing factors in perception. Journal of Personality, 1947, 15, 300-308.

- Burkhart, B. R., Christian, W. L., & Gynther, M. D. Item subtlety and faking on the MMPI: A paradoxical relationship. Journal of Personality Assessment, 1978, 42, 76-80.
- Burkhart, B. R., Gynther, M. D., & Christian, W. L. Psychological mindedness, intelligence, and item subtlety endorsement patterns on the MMPI. Journal of Clinical Psychology, 1978, 34, 76-79.
- Buros, O. K. (Ed.). The eighth mental measurements yearbook. Highland Park, N. J.: Gryphon Press, 1978.
- Campbell, D. T. A typology of tests, projective and otherwise. Journal of Consulting Psychology, 1957, 21, 207-210.
- Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. Handbook for the Sixteen Personality Factor Questionnaire (16 PF). Champaign, Ill.: Institute for Ability Testing, 1970.
- Christian, W. L., Burkhardt, B. R., & Gynther, M. D. Subtle-obvious ratings of MMPI items: New interest in an old concept. Journal of Consulting and Clinical Psychology, 1978, 46, 1178-1186.
- Cofer, C. N., Chance, J., & Judson, A. J. A study of malin-gering on the Minnesota Multiphasic Personality Inventory. Journal of Psychology, 1949, 27, 491-499.
- Cohn, T. S. Is the F scale indirect? Journal of Abnormal and Social Psychology, 1952, 47, 732.

- Cooke, G., & Robey, A. The MMPI: A case study in dissimulation. Journal of Consulting and Clinical Psychology, 1971, 36, 355-359.
- Cronbach, L. J. Response set and test validity. Educational and Psychological Measurement, 1946, 6, 475-494.
- Cronbach, L. J. Further evidence on response sets and test validity. Educational and Psychological Measurement, 1950, 10, 3-31.
- Cronbach, L. J. Essentials of psychological testing (3rd ed.). New York: Harper & Row, 1970.
- Crowne, D. P., & Marlowe, D. A new scale of social desirability independent of psychopathology. Journal of Consulting Psychology, 1960, 24, 349-354.
- Dahlstrom, W. G., & Welsh, G. S. An MMPI handbook. Minneapolis: University of Minnesota Press, 1960.
- Dahlstrom, W. G., Welsh, G. S., & Dahlstrom, L. E. An MMPI handbook, volume II (Rev. ed.). Minneapolis: University of Minnesota Press, 1975.
- Davids, A. Comparison of three methods of personality assessment: Direct, indirect, and projective. Journal of Personality, 1955, 23, 423-440.
- Drasgow, J., & Barnette, W. L., Jr. F-K in a motivated group. Journal of Consulting Psychology, 1957, 21, 399-401.

Duff, F. L. Item subtlety in personality inventory scales.

Journal of Consulting Psychology, 1965, 29, 565-570.

Edwards, A. L. The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. Journal of Applied Psychology, 1953, 37, 90-93.

Edwards, A. L. The social desirability variable in personality assessment and research. New York: Dryden Press, 1957.

Edwards, A. L. Social desirability or acquiescence in the MMPI? A case study with the SD scale. Journal of Abnormal and Social Psychology, 1961, 63, 351-359.

Edwards, A. L., & Diers, C. J. Social desirability and the factorial interpretation of the MMPI. Educational and Psychological Measurement, 1962, 22, 501-509.

Edwards, A. L., & Walsh, J. A. The relationship between the intensity of the social desirability keying of a scale and the correlation of the scale with Edwards' SD scale and the first factor loading of the scale. Journal of Clinical Psychology, 1963, 19, 200-203.

Erickson, C. W., & Browne, C. T. An experimental and theoretical analysis of perceptual defense. Journal of Abnormal and Social Psychology, 1956, 52, 224-230.

- Exner, J. E., Jr., McDowell, E., Pabst, J., Stackman, W., & Kirk, L. On the detection of willful falsification in the MMPI. Journal of Consulting Psychology, 1963, 27, 91-94.
- Fink, A. M., & Butcher, J. N. Reducing objections to personality inventories with special instructions. Educational and Psychological Measurement, 1972, 32, 631-639.
- Fiske, D. W. An intensive study of variability scores. Educational and Psychological Measurement, 1957, 17, 453-465. (a)
- Fiske, D. W. The constraints on intra-individual variability in test responses. Educational and Psychological Measurement, 1957, 17, 317-337. (b)
- Fiske, D. W. Measuring the concepts of personality. Chicago: Adline Publishing Company, 1971.
- Fiske, D. W., & Butler, J. M. The experimental conditions for measuring individual differences. Educational and Psychological Measurement, 1963, 23, 249-265.
- Fordyce, W. E. Social desirability in the MMPI. Journal of Consulting Psychology, 1956, 20, 171-175.
- Fricke, B. G. Subtle and obvious test items and response set. Journal of Consulting Psychology, 1957, 21, 250-252.
- Gloye, E. E., & Zimmerman, I. L. MMPI item changes by college students under ideal-self response set. Journal of Projective Techniques and Personality Assessment, 1967, 31, 63-69.

- Goldberg, L. R. A model of item ambiguity in personality assessment. Educational and Psychological Measurement, 1963, 23, 467-492.
- Goldberg, L. R., & Slovic, P. Importance of test item content: An analysis of a corollary of the deviation hypothesis. Journal of Counseling Psychology, 1967, 14, 462-472.
- Goldfried, M. R. Behavioral assessment. In I. B. Weiner (Ed.), Clinical methods in psychology. New York: John Wiley and Sons, 1976.
- Gough, H. G. Simulated patterns on the Minnesota Multiphasic Personality Inventory. Journal of Abnormal and Social Psychology, 1947, 42, 215-225.
- Gough, H. G. The F minus K dissimulation index for the Minnesota Multiphasic Personality Inventory. Journal of Consulting Psychology, 1950, 14, 408-413.
- Grayson, H. M., & Olinger, L. B. Simulation of "normalcy" by psychiatric patients on the MMPI. Journal of Consulting Psychology, 1957, 21, 73-77.
- Gynther, M. D., Burkhart, B. R., & Hovanitz, C. Do face-valid items have more predictive validity than subtle items? The case of the MMPI Pd scale. Journal of Consulting and Clinical Psychology, 1979, 47, 295-300.

- Hamsher, J. H. Disguise of purpose in psychological testing. Proceedings of the 77th Annual Convention of the American Psychological Association, 1969, 4, 141-142. (Summary)
- Hanley, C. Deriving a measure of test-taking defensiveness. Journal of Consulting Psychology, 1957, 21, 391-397.
- Hanley, C. Social desirability and response bias in the MMPI. Journal of Consulting Psychology, 1961, 25, 13-20.
- Harris, J. G., Jr., & Baxter, J. C. Ambiguity in the MMPI. Journal of Consulting Psychology, 1965, 29, 112-118.
- Harvey, M. A., & Sipprelle, C. N. Demand characteristic effects on the subtle and obvious subscales of the MMPI. Journal of Personality Assessment, 1976, 40, 539-544.
- Heron, A. The effects of real-life motivation on the questionnaire response. Journal of Applied Psychology, 1956, 40, 65-68.
- Hiner, D. L., Ogren, D. J., & Baxter, J. C. Ideal-self responding on the MMPI. Journal of Projective Techniques and Personality Assessment, 1969, 33, 389-396.
- Hoffman, E. L. The effect of instructions to be "accurate" on responses to the MMPI (Doctoral dissertation, University of Illinois, 1968). Dissertation Abstracts, 1968, 29, 757B. (University Microfilms No. 68-12, 130)
- Holden, R. R., & Jackson, D. N. Item subtlety and face validity in personality assessment. Journal of Consulting and Clinical Psychology, 1979, 47, 459-468.

- Horlick, R. S. The discriminant value of Minnesota Multiphasic Personality Inventory items in personality disorders. Journal of Clinical Psychology, 1955, 11, 362-365.
- Hunt, H. F. The effects of deliberate deception on Minnesota Multiphasic Personality Inventory performance. Journal of Consulting Psychology, 1948, 12, 396-402.
- Hunt, H. F. The differentiation of malingering, dissimulation, and pathology. In M. Hammer, K. Salzinger & S. Sutton (Eds.), Psychopathology: Contributions from the social, behavioral, and biological sciences. New York: Wiley and Sons, 1973.
- Jackson, D. N. The dynamics of structured personality tests: 1971. Psychological Review, 1971, 78, 229-248.
- Jackson, D. N., & Messick, S. Content and style in personality assessment. Psychological Bulletin, 1958, 55, 243-252.
- Jackson, D. N., & Messick, S. A distinction between judgments of frequency and of desirability as determinants of response. Educational and Psychological Measurement, 1969, 29, 273-293.
- Kaufman, P. Changes in the Minnesota Multiphasic Personality Inventory as a function of psychiatric therapy. Journal of Clinical Psychology, 1950, 14, 458-464.

- Kimber, J. A. M. The insight of college students into the items on a personality test. Educational and Psychological Measurement, 1947, 7, 411-420.
- Kroger, R. O. Faking in interest measurement: A social psychological perspective. Measurement and Evaluation in Guidance, 1974, 7(2), 130-134.
- Kroger, R. O., & Turnbull, W. Effects of role demands and test-cue properties on personality test performance: Replication and extension. Journal of Consulting and Clinical Psychology, 1970, 35, 381-387.
- Lanyon, R. I. Simulation of normal and psychopathic MMPI personality patterns. Journal of Consulting Psychology, 1967, 31, 94-97.
- Lawton, M. P. Deliberate faking on the Psychopathic Deviate scale of the MMPI. Journal of Clinical Psychology, 1963, 19, 327-330.
- Lawton, M. P., & Kleban, M. H. Prisoners' faking on the MMPI. Journal of Clinical Psychology, 1965, 21, 269-271.
- Levitt, H., & Rice, D. Sex and verbal proficiency on the simulation of normalcy procedure. Journal of Clinical Psychology, 1968, 24, 440-443.
- Lingoes, J. C. MMPI factors of the Harris and Wiener subscales. Journal of Consulting Psychology, 1960, 24, 74-83.

- MacLean, A. G., Tait, A. T., & Catterall, C. D. The F minus K index on the MMPI. Journal of Applied Psychology, 1953, 37, 315-316.
- Manson, M. P. A psychometric analysis of the psychopathic characteristics of alcoholics. Journal of Consulting Psychology, 1949, 13, 111-118.
- Match, J., & Wiggins, N. Individual viewpoints of social desirability related to faking good and desirability estimation. Educational and Psychological Measurement, 1974, 34, 591-605.
- Mathews, A., & Wertheimer, M. A "pure" measure of perceptual defense uncontaminated by response suppression. Journal of Abnormal and Social Psychology, 1958, 57, 373-376.
- Mayo, G. D., & Guttman, I. Faking in a vocational classification situation. Journal of Applied Psychology, 1959, 43, 117-121.
- McCall, R. J. Face validity in the D scale of the MMPI. Journal of Clinical Psychology, 1958, 14, 77-80.
- McKinley, J. C., & Hathaway, S. R. A multiphasic personality schedule (Minnesota): II. A differential study of hypochondriasis. Journal of Psychology, 1940, 10, 255-268.
- McKinley, J. C., Hathaway, S. R., & Meehl, P. E. The Minnesota Multiphasic Personality Inventory: VI. The K scale. Journal of Consulting Psychology, 1948, 12, 20-31.

- Meehl, P. E. An investigation of a general normality or control factor in personality testing. Psychological Monographs, 1945, 59(4, Whole No. 274). (a)
- Meehl, P. E. The dynamics of "structured" personality tests. Journal of Clinical Psychology, 1945, 1, 296-303. (b)
- Meehl, P. E., & Hathaway, S. R. The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. Journal of Applied Psychology, 1946, 30, 525-564.
- Merbaum, M. Simulation of normal MMPI profiles by repressors and sensitizers. Journal of Consulting and Clinical Psychology, 1972, 39, 171.
- Messick, S., & Jackson, D. N. Acquiescence and the factorial interpretation of the MMPI. Psychological Bulletin, 1961, 58, 299-304. (a)
- Messick, S., & Jackson, D. N. Desirability scale values and dispersions for MMPI items. Psychological Reports, 1961, 8, 409-414. (b)
- Milgram, N. A. Cognitive and empathetic factors in role-taking by schizophrenic and brain damaged patients. Journal of Abnormal and Social Psychology, 1960, 59, 219-224.
- Millman, J., Bishop, C. H., & Ebel, R. An analysis of test-wiseness. Educational and Psychological Measurement, 1965, 25, 717-726.

- Morf, M. E., & Jackson, D. N. An analysis of two response styles: True responding and item endorsement. Educational and Psychological Measurement, 1972, 32, 329-353.
- Noll, V. H. Simulation by college students of a prescribed pattern on a personality scale. Educational and Psychological Measurement, 1951, 11, 478-488.
- Norman, R. P. Extreme response tendency as a function of emotional adjustment and stimulus ambiguity. Journal of Consulting and Clinical Psychology, 1969, 33, 406-410.
- Norman, W. T. Relative importance of test item content. Journal of Consulting Psychology, 1963, 27, 175-178.
- Orne, M. T. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. American Psychologist, 1962, 17, 776-783.
- Osgood, C. E. The nature and measurement of meaning. Psychological Bulletin, 1952, 49, 197-237.
- Osgood, C. E. Studies on the generality of affective meaning systems. American Psychologist, 1962, 17, 10-28.
- Parsons, O. A., Yourshaw, S., & Borstelmann, L. Self-ideal-self discrepancies on the MMPI: Consistencies over time and geographic region. Journal of Counseling Psychology, 1968, 15, 160-166.

- Payne, F. D., & Wiggins, J. S. MMPI profile types and the self-report of psychiatric patients. Journal of Abnormal Psychology, 1972, 79, 1-8.
- Rapaport, G. M. "Ideal self" instructions, MMPI profile changes, and the prediction of clinical improvement. Journal of Consulting Psychology, 1958, 22, 459-463.
- Rorer, L. G. The great response style myth. Psychological Bulletin, 1965, 63, 129-156.
- Rosen, E. Self-appraisal, personal desirability, and perceived social desirability of personality traits. Journal of Abnormal and Social Psychology, 1956, 52, 151-158.
- Sacks, J. M., & Kirtley, D. D. Some personality characteristics related to response to subtle and obvious items on the MMPI. Journal of Consulting and Clinical Psychology, 1972, 38, 66-69.
- Seeman, W. "Subtlety" in structured personality tests. Journal of Consulting Psychology, 1952, 16, 278-283.
- Seeman, W. Concept of "subtlety" in structured psychiatric and personality tests: An experimental approach. Journal of Abnormal and Social Psychology, 1953, 48, 239-247.
- Sheldon, M. S. Conditions affecting the fakability of teacher-selection inventories. Educational and Psychological Measurement, 1959, 19, 207-219.

- Siller, J., & Chipman, A. Response set paralysis: Implications for measurement and control. Journal of Consulting Psychology, 1963, 27, 432-438.
- Singer, W. B., & Young, P. T. Studies in affective reaction: III. The specificity of affective reactions. Journal of General Psychology, 1941, 24, 327-341.
- Stone, L. A. Subtle and obvious response on the MMPI as a function of acquiescence response style. Psychological Reports, 1965, 16, 803-804.
- Stricker, L. J. "Test-wiseness" on personality scales. Journal of Applied Psychology Monograph, 1969, 53(3, Pt. 2).
- Taylor, J. B., Carithers, M., & Coyne, L. MMPI performance, response set, and the "self concept hypothesis." Journal of Consulting and Clinical Psychology, 1976, 44, 351-362.
- Tracy, R. L., & Fiske, D. W. Does information about personality scales affect responses and attitudes? Educational and Psychological Measurement, 1974, 34, 617-629.
- Vesprani, G. J., & Seeman, W. MMPI X and zero items in a psychiatric outpatient group. Journal of Personality Assessment, 1974, 38, 61-64.
- Voas, R. B. A procedure for reducing the effects of slanting questionnaire responses toward social acceptability. Educational and Psychological Measurement, 1958, 18, 337-345.

- Wales, B., & Seeman, W. A new method for detecting the fake good response set on the MMPI. Journal of Clinical Psychology, 1968, 24, 211-216.
- Wales, B., & Seeman, W. What do MMPI zero items really measure: An experimental investigation. Journal of Clinical Psychology, 1969, 25, 420-424.
- Wales, B., & Seeman, W. Instructional set and MMPI items. Journal of Personality Assessment, 1972, 36, 282-286.
- Wiener, D. N. Selecting salesman with the subtle-obvious keys for the MMPI. American Psychologist, 1948, 3, 364.
(Abstract) (a)
- Wiener, D. N. Subtle and obvious keys for the Minnesota Multiphasic Personality Inventory. Journal of Consulting Psychology, 1948, 12, 164-170. (b)
- Wiener, D. N. A control factor in social adjustment. Journal of Abnormal and Social Psychology, 1951, 46, 3-8.
- Wiggins, J. S. Interrelationships among MMPI measures of dissimulation under standard and social desirability instructions. Journal of Consulting Psychology, 1959, 23, 419-427.
- Wiggins, J. S. Strategic, method, and stylistic variance in the MMPI. Psychological Bulletin, 1962, 59, 224-242.
- Wiggins, J. S. Social desirability estimation and "faking good" well. Educational and Psychological Measurement, 1966, 26, 329-341.

- Wilcox, R., & Krasnoff, A. Influence of test-taking attitudes on personality inventory scores. Journal of Consulting Psychology, 1967, 31, 188-194.
- Windle, C. Further studies of test-retest effect on personality questionnaires. Educational and Psychological Measurement, 1955, 15, 246-254.
- Winne, J. F. A scale of neuroticism: An adaptation of the Minnesota Multiphasic Personality Inventory. Journal of Clinical Psychology, 1951, 7, 117-122.
- Zimmerman, I. L., & Gloye, E. E. A definition and analysis of the "ego ideal" based upon differential response sets to the MMPI. American Psychologist, 1959, 14, 364.
- (Abstract)