University of Montana

ScholarWorks at University of Montana

Graduate Student Theses, Dissertations, &
Professional Papers

Graduate School

1989

Investigation of response bias in the Chapman scales

Bryan David Peltier

The University of Montana

Follow this and additional works at: <https://scholarworks.umt.edu/etd>

Let us know how access to this document benefits you.

Recommended Citation

Peltier, Bryan David, "Investigation of response bias in the Chapman scales" (1989). *Graduate Student Theses, Dissertations, & Professional Papers*. 4994.
<https://scholarworks.umt.edu/etd/4994>

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.

COPYRIGHT ACT OF 1976

THIS IS AN UNPUBLISHED MANUSCRIPT IN WHICH COPYRIGHT
SUBSISTS. ANY FURTHER REPRINTING OF ITS CONTENTS MUST BE
APPROVED BY THE AUTHOR.

MANSFIELD LIBRARY
UNIVERSITY OF MONTANA
DATE: 1989

An Investigation of Response Bias in the Chapman Scales

by

Bryan David Peltier
B.S., University of Nevada - Reno, 1985

Presented in Partial Fulfillment of the Requirements
for the degree of
Masters of Arts
University of Montana, 1989

Approved by:

James A. Walsh
Chairperson, Board of Examiners

RC Murray
Dean, Graduate School

October 11, 1989
Date

UMI Number: EP40458

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.

UMI EP40458

Published by ProQuest LLC (2014). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

11-29-89
Feltier, Bryan D., M.A., June 1989

Clinical Psychology

An investigation of response bias in the Chapman Scales (76 pp.)

Director: James A. Walsh, Ph.D. *James A. Walsh*

This study examined response set problems associated with four of the Chapman scales, which reportedly measure personality traits of individuals considered to be prone to psychosis. Evidence suggested that interpretations of the Chapman scales were confounded by response set biases, and that the Chapman scales have been simply measuring, to a considerable extent, the tendency of subjects to give socially desirable responses in self-description. Revised normative data for the Edwards and Marlowe-Crowne social desirability scales indicated that social desirability norms are less rigorous than they were twenty years ago. Multiple regression and correlation techniques indicated socially desirable and acquiescent response styles may account for 25% of the total variance in the Perceptual Aberration, Magical Ideation, and Impulsive Nonconformity Scales. However, no reliable response set predictors were identified for the Physical Anhedonia Scale. Results demonstrated that normal subjects were able to identify the keyed responses on each of these selected Chapman scales, and successfully fake their scores.

Table of Contents

Introduction.	i
Method.	34
Results	36
Discussion.	50
References.	57
Appendix A: Edwards Social Desirability Scale	61
Appendix B: Marlowe-Crowne SD Scale	63
Appendix C: Jackson's Infrequency Scale	65
Appendix D: Welsh's R Scale	66
Appendix E: Faking Instructions	68
Appendix F: Chapman Scales	69
Appendix G: Correlations	76

List of Tables

Table 1.	Correlations between the SD scale and the MMPI scales.	.11
Table 2.	Current and past normative data for the Edwards and Marlowe-Crowne social desirability scales, Welsh's R Scale, and Jackson's Infrequency Scale	36
Table 3.	Correlations between the Edwards and Marlowe-Crowne social desirability scales, Welsh's R Scale, and Jackson's Infrequency Scale.	.37
Table 4.	Means and Standard Deviations for the Perceptual Aberration Scale, the Magical Ideation Scale, the Physical Anhedonia Scale, and the Impulsive Nonconformity Scale under Standard Instructions.	38
Table 5.	Average Correlations between the Perceptual Aberration Scale, the Magical Ideation Scale, the Physical Anhedonia Scale, and the Impulsive Nonconformity Scale	40
Table 6.	Correlations between the Chapman Scales and the Response Set Measures	40
Table 7.	Prediction of Chapman Scales from Response Set Measures - Stepwise Regression Analysis	42
Table 8.	Double-Split Cross Validation of Stepwise Regression of the Chapman Scales on the Response Set Measures.	44
Table 9.	Means and Standard Deviations for the Chapman Scales under Faking Instructions.	45
Table 10.	Means and Standard Deviations for Per-Mags on the Chapman Scales under Faking Instructions.	48
Table 11.	Correlations between the Chapman Scales and the Response Set Measures for Per-Mags.	49

An Investigation of Response Bias in the Chapman Scales

It has long been observed that scores on personality questionnaires are influenced by factors other than the manifest content of the items. Simply because an individual responds "false" to the item, "I often think about things too bad to talk about," it cannot be determined whether this person's thoughts would not be upsetting to others, whether bad thoughts are experienced but the individual refuses for any number of reasons to divulge that information, or whether the individual is generally disposed to disagree with most personality test items regardless of their content. Conversely, a "true" response to this item would not rule out possibility that the respondent was displaying a tendency to agree with most personality statements irrespective of their content.

The omnipresence of this problem - that responses to personality test items cannot always be taken at face value - has been a considerable source of frustration to psychologists in their quests for accurate assessments, developing successful treatment applications, and further understanding of the human experience. While psychologists are advised to attempt to elicit the client's or subject's full cooperation before beginning a test, there are additional factors operating that must be considered when interpreting responses to items on a personality questionnaire.

What, then, are the sources of variance in personality

tests? In searching for the answer to this question, the concept of response set was introduced, incorporating the idea that subjects bring sets, or habitual response preferences, to the testing situation. These include the tendency to indiscriminantly agree with personality statements, and the tendency to describe oneself in favorable, socially desirable terms, either as a result of a general learning set, or in order to achieve the approval of others.

The development of the Minnesota Multiphasic Personality Inventory capitalized on the growing awareness of the response bias problem. In the construction of the MMPI, three special validity scales (F, L, and K) were incorporated to detect and, in the case of the K scale, to suppress, defensive or "fake good" tendencies. The F, L, and K validity indicators, as well as other indices, attempted to detect any inappropriate response sets that might be utilized by the test-taker. (These scales will be described in detail below). For example, a person with limited reading ability may complete a given questionnaire very quickly, which would arouse legitimate concern about the authenticity of the item endorsements. Another possibility is that a person may be unable or unwilling to take the test, but instead of refusing directly to complete it, they may endorse all items "true" or "false," or alternate "true" and

"false" responses. The validity scales of the MMPI attempt to identify and correct these response set problems. These problems must be adequately addressed with any personality inventory before reasonable interpretations of the scales can be made.

Problems associated with response biases are particularly devastating to the validity of personality scales that are rationally constructed (Edwards, 1970). Before attempting to measure individual differences with respect to some personality construct, it is necessary to demonstrate convergent and discriminant validity for the scale being used. Yet, for rational scales there seldom exists an external criterion that can be regarded as accurately measuring the trait which the rational scale is designed to measure. Instead, rational scales are ordinarily validated by hypothesis testing and experimentation.

Therefore, it is extremely important to develop personality scales that are relatively free from response biases to ensure that the scale is not simply measuring personal sets to respond to content-irrelevant aspects of the test. The major response sets that interfere with the interpretation of personality scales are described below.

Edwards Social Desirability Response Set

Edwards discussed the acquisition of the social desirability (SD) response set in terms of social rein-

forcement for "learning cultural norms of what is desirable and undesirable in the way of personality traits and characteristics" (Edwards, 1970, p. 224). In every society an effort is made to teach children those standards of conduct that society considers "good" and "bad," or socially desirable and undesirable. Children are rewarded for acts that are considered to be desirable, and punished for those that are considered undesirable. It seems reasonable to assume that by early adolescence, if not before, almost all children have acquired a fairly good understanding of what is socially desirable and undesirable within the context of their particular culture.

Operationalizing this view, Edwards viewed an SD response as a "true" response to an item where a panel of subjects has judged it to be socially desirable to say "true," or a "false" response to an item to which it has been judged socially undesirable (SUD) to say "true."

If one considers the population of statements that might be used in describing personality, each statement can be characterized in terms of its position on a single dimension of social desirability-undesirability, or, more briefly, social desirability. It is possible to obtain a scale value for any personality statement such that the scale value indicates the position of the statement on the social desirability continuum, relative to other statements.

Regardless of the multidimensional nature of personality statements with respect to content, it is nonetheless possible to describe each one in terms of its position on the social desirability dimension. According to Edwards, the social desirability continuum appears to be the most important single dimension on which to locate personality statements by virtue of the fact that if we know the position of a statement on it, we can then predict, with a high degree of accuracy, the proportion of individuals who will say, in self-description, that the statement describes them (Edwards, 1957).

Thus, the social desirability scale value (SDSV) of a personality statement refers to its location on the continuum of social desirability as determined by one of the various psychological scaling methods. High SDSVs represent statements that are judged as highly socially desirable, and low SDSVs represent statements that are judged as socially undesirable, with reference to an arbitrary origin of average social desirability. The relationship between the social desirability scale value of a personality statement and the probability that this statement will be endorsed by subjects when they are asked to describe their own personalities in terms of the statement will be described next.

Suppose that a given set of personality statements has

been rated for social desirability, and that the SDSVs of the statements are available. If this same set of statements is given to another independent group of individuals, and they are asked to describe themselves by responding "true" or "false" to each statement, then it is possible to obtain the proportion of "true" responses, that is, $P(T)$ given to each statement. $P(T)$ is called the probability of item endorsement. Then, for each item, we would have a value of $P(T)$ and a value of SDSV.

For example, a set of 140 personality statements was judged for social desirability by a group of 152 judges (Edwards, 1953). Scale values for the statements were obtained by the method of successive intervals. The statements were printed in inventory form, and then administered to a new group of 152 students. One of the statements in the inventory was "I like to be loyal to my friends." This statement had an extremely high social desirability scale value, and was endorsed by 98 percent of the students. Another statement in the inventory was "I like to avoid responsibilities and obligations." This statement had an extremely low social desirability scale value, and was endorsed by only 6 percent of the students. Figure 1 shows the relationship between $P(T)$ and SDSV for the 140 personality statements.

If $P(T)$ is plotted against SDSV for any random or

Figure 1. Probability of endorsement of a personality statement as related to the social desirability scale value of the statement. Statements are from the Personal Preference Schedule. Source: Edwards (1953).

representative set of statements, it will be found that $P(T)$ increases linearly with SDSV. In other words, statements with low or socially undesirable scale values have a low probability of being answered true, and statements with high or socially desirable scale values have a high probability of being answered true. Various studies have shown that $P(T)$ is linearly related to SDSV (Cowen & Tongas, 1959; Edwards & Walsh, 1963a; Hanley, 1956; Kenny, 1956; Taylor, 1959). The linear relationship is found not only in college and adult samples, but also when young children are asked to describe themselves. For example, the correlation between $P(T)$ and SDSV has been reported by Cruse (1966) to increase from 0.61 for children age 3, to 0.71 for children age 4, to .74 for children age 5, and to 0.88 for children age 6. Additionally, there is evidence that children's judgments of the social desirability of personality statements are highly correlated with the SDSVs of the statements based on the judgements of adults. Walsh, Tomlinson-Keasey, and Klieger (1974) demonstrated that children in the first grade can give ratings of SDSVs for items that correlate as high as +.92 with rating by college students.

However, it would be erroneous to conclude that simply because an individual has learned the cultural norms of social desirability, he or she will necessarily behave in accordance with them, or describe himself or herself in

terms of them. Yet, individual differences in rates of SD responding represent a reliable personality trait, and rational scales can be developed to measure this trait without regard to the content of the items in the scale other than as this content determines the SDSVs of the items. Thus, the social desirability response set has been shown to be related to responses to a wide variety of objective personality, attitude, and interest scales.

The Edwards Social Desirability Scale.

A procedure frequently used in constructing attitude scales is to have a group of subjects judge the degree of favorableness or unfavorableness of each of a number of attitude statements relating to some psychological construct. Upon the basis of these judgements a scale value is obtained for each statement by one of the psychological scaling methods. The scale value of a statement is taken as an indicator of the location of the statement on a psychological continuum ranging from highly unfavorable, through neutral, to highly favorable. There is adequate and ample evidence to indicate that the scale values thus obtained are highly reliable and relatively independent of the attitudes of the judges themselves toward the psychological construct. (Edwards, 1957).

In addition, if a scale was designed to measure a personality trait that was itself considered socially desir-

able, then one would expect to find the SD key for the items in the scale to be very similar to the trait key. If this scale was scored by each key, the correlation between the two sets of scores should be quite high and positive. Similarly, if a high score on a given personality scale indicates the presence of a trait that would itself be judged as socially undesirable, one would expect the scoring keys for SD and the trait to be reversed, thus producing a high negative correlation between the SD and trait scores.

An initial set of 150 items was obtained from the F, L, K, and Anxiety scales of the MMPI (Edwards, 1957). These were submitted to a group of 10 judges who were asked to give socially desirable responses to the items. For 79 of the 150 items perfect agreement was found that a particular response was the judged socially desirable response. The items were later analyzed, and a set of 39 items were selected that showed the greatest differentiation between a high and low group in terms of total SD scores on the 79 item scale. This 39 item SD scale has since been used in a variety of investigations (Merrill & Heathers, 1956; Welsh, 1956; Edwards, 1961; Messick & Jackson, 1961a; Edwards, Diers, & Walker, 1962; Edwards & Walsh, 1963b, 1964; Block, 1965).

If the SD scale provides a measure of the tendency of

subjects to give socially desirable responses to statements in self-description, then the correlations of scores on this scale with other personality scales, given under standard instructions, should indicate something of the extent to which the social desirability variable is operating. Cronbach and Meehl (1955), in a discussion of the notion of construct validity in psychological tests, pointed out that if two tests are presumed to measure the same construct, a correlation between them can be predicted. Table 1 gives the correlations between the SD scale and various clinical and validity scales of the MMPI as obtained by Merrill and Heathers (1956). The group tested by these investigators consisted of 155 counseling center clients. Note that the

Table 1.

Correlations between the SD scale and the MMPI scales.

MMPI Scale	Edwards SD Scale
L	.14
F	-.63
K	.81
Hs	-.52
D	-.61
Hy	.08
Pd	-.50
Mf	-.36
Pa	-.09
Pt	-.85
Sc	-.77
Ma	-.13

Source: Merrill and Heathers (1956), N = 155.

correlation between the SD scale and the K scale is very high.

The K scale of the MMPI consists of 30 items that were empirically selected to assist in identifying persons who displayed significant pathology, yet had profiles within the normal range. The objective of the K scale was to provide a correction for scores on certain diagnostic scales. The content of the items on the K scale is heterogeneous, and covers self-control, and family and interpersonal relationships. Examples of the K scale items with the deviant answer indicated in parentheses are as follows:

"I like to let people know where I stand on things."
(false)

"I have very few quarrels with members of my family."
(true)

"People often disappoint me." (false)

The K-corrected profile of the MMPI was developed by determining the proportion of K that, when added to the raw score on the clinical scales, would maximize the discrimination between the normative groups and the criterions groups. The K scale is considered a measure of personality integration and healthy adjustment, with high scores reflecting more positive adjustment. Very high scores on the K scale are taken to indicate people who are consistently trying to maintain a facade of adequacy and control,

and are admitting no problems or weaknesses ("faking good"). Presumably, individuals with high scores on the K scale have been somewhat defensive in answering the items in the MMPI, and consequently have failed to obtain as high a score on some of the diagnostic scales as they would have obtained if they had been less defensive. Low scores suggest that these persons are greatly exaggerating their problems to create an unfavorable impression ("faking bad").

Adjustment and social desirability may not be equivalent concepts. However, they are quite similar in meaning. One should expect, therefore, to find scores on the K scale highly correlated with those on the SD scale. Table 1 indicates this is the case.

However, if a statement has a high SDSV, and if a subject endorses it in self-description, interpretation of this response is confounded unless it can be demonstrated that SD response biases are not involved. On an anxiety scale, for instance, each of the items would be keyed for the response that supposedly indicates the presence of this trait. Suppose that the items in the anxiety scale are rated for social desirability and the SDSV of each of the items is obtained. If upon examining the trait responses to the items, it is then discovered that each trait response is also an SUD response, we would say that the trait and the SUD responses to the items are completely confounded.

Scores on the anxiety scale may simply reflect the tendency to give SD responses in self-description. Then it would also be possible to describe a low scoring group on the anxiety scale as those who desire to make a good impression on others, and a high scoring group as those who are less interested in what others think of them.

Factor analysis and other studies of the internal structure of personality tests give evidence that this type of confounding of trait and SD responses occurs frequently. These studies have revealed very few factors, often not more than two, which have been demonstrated to be associated with the response styles of social desirability and acquiescence. (Jackson, 1960; Jackson & Messick, 1961, 1962a, 1962b; Messick & Jackson, 1961b; Edwards, 1963; Wiggins, 1962). It has been shown in principal-component factor analyses of MMPI scales that the Edwards SD scale has a high loading at one pole of the first MMPI factor often described as a psychotic factor. MMPI scales having a large proportion of items keyed for SD responses also tend to have relatively high loadings at the same pole of the first factor as the Edwards SD scale. MMPI scales having a large proportion of items keyed for SUD responses tend to have relatively high loadings at the opposite pole of the first factor. Thus, the proportion of items keyed for SD responses in the MMPI scales has been shown to be substantially correlated with

the first factor loadings of the scales (Edwards, 1970). The relationship between the various clinical scales of the MMPI and the SD scale suggests, according to Fordyce (1956, p. 174), "that mental illness, in a generalized sense of the term, may be characterized as involving behavior which is socially disapproved."

Marlowe-Crowne Social Desirability

Social approval scales of the Marlowe-Crowne type have been shown to have low-to-moderate correlations with the kind of SD scales described by Edwards. The fact that both kinds of scales have been referred to as social desirability scales frequently leads to confusion.

Imagine a class of personality statements that is defined by two major attributes: First, they are culturally sanctioned things to say about oneself; and second, they are probably untrue of most people. A balanced scale composed half of culturally acceptable but probably untrue statements, and half of true but undesirable statements constitutes the model for the Marlowe-Crowne Social Desirability (M-C SD) scale (Crowne & Marlowe, 1964).

The Marlowe - Crowne Social Desirability Scale.

The Marlowe-Crowne Social Desirability Scale was constructed in the following manner. Items were selected that met the criterion of cultural approval and yet were untrue of virtually all people, and that had minimal

pathological implications. A set of 50 items meeting these criteria was submitted to ten judges for social desirability ratings. The judges were instructed to score the items in the socially desirable direction using true and false response categories. Unanimous agreement was obtained on 36 items and 90% on 11 more. These 47 items constituted the preliminary form of the scale.

The preliminary version of the scale was then administered to 76 students in two introductory psychology classes, and an item analysis performed. There were 33 items which discriminated between high and low total scores at the .05 level or better. Of the 33 items, 18 are keyed true and 15 false. These 33 items constitute the final form of the M-C SD scale.

The M-C SD scale is very similar to the Lie scale of the MMPI. The Lie (L) scale of the MMPI includes 15 items that were selected on a rational basis to identify persons who are deliberately trying to avoid answering the MMPI frankly and honestly. The L scale also assesses attitudes and practices that are culturally commendable, but few individuals could in honesty, consistently give the keyed responses. Content areas within the L scale include denial of minor, personal dishonesties and denial of aggression, bad thoughts, and weakness of character. Examples of the L scale items with the deviant answer indicated in parentheses

are as follows:

"At times I feel like swearing." (false)

"I get angry sometimes." (false)

"Sometimes when I am not feeling well I am cross."
(false)

While there are some individuals who could give the false response to several items on the scale without necessarily being dishonest, almost all individuals who take the MMPI obtain low scores on the L scale. Thus, higher scores are regarded as revealing something about an individual's tendency not to be frank.

Since "false" is the deviant answer to all L scale items, the L scale is extremely susceptible to unsophisticated deviant test-taking sets, such as the set to blatantly seek social approval, or the set to answer all items "true." More sophisticated deviant response sets may go undetected by the L scale. Inspection of the L scale items reveals that it is readily apparent which responses are the deviant ones. Numerous studies have shown that the L scale does not detect sophisticated persons who were given instructions to falsify their answers to the MMPI (Dahlstrom et al., 1972; Vincent, Linsz, & Green, 1966). These persons apparently realized that it would be unconvincing to give the keyed responses to L scale items. Thus, the L scale can be construed as a measure of psychological sophistication with

high scores indicating a lack of such sophistication. The person's education level and socioeconomic class must be kept in mind when interpreting the L scale.

The interpretation of socially desirable responding on the M-C SD scale is somewhat different. It is believed that people conform to social stereotypes of what is good to acknowledge about oneself in order to achieve approval from others. Individuals who depict themselves in very favorable terms on the M-C SD scale are considered to be displaying a strong need for social approval, and thus have a social desirability response set in this more specific sense (Crowne & Marlowe, 1964).

Acquiescence

Acquiescence, the tendency to mark true, indicate agreement, or otherwise respond uncritically to test items by personal endorsement, has been established as a major response determinant in the measurement of personality variables (Bass, 1955; Chapman & Campbell, 1957; Jackson & Messick, 1958). Couch and Keniston (1960) have in fact shown that "yea-saying" is widely general over different items and tests, and they suggest that this response set stems from "a central personality syndrome" (p.173).

It has long been recognized by those who have constructed personality scales that approximately half of the statements in the scale should be keyed in the opposite

direction from the other half. In this way, one hopes to avoid the confounding of scores with the tendency to respond consistently "true" or "false" regardless of the content of the items.

If a majority of the items in an inventory are keyed "true," for example, a high score may be measuring not only the variable of interest, but also the tendency of the subject to acquiesce. Similarly, if a majority of the items in the scale is keyed "false," then high scores may also be measuring the tendency of the subject to dissent, that is, to generally disagree with the items in the scale. Thus, scores on two scales in which all of the items are keyed "true" might be expected to have some degree of positive correlation simply because of individual differences in acquiescent tendencies. Therefore, scores on personality scales are regarded as being susceptible to the influence of acquiescent tendencies to the degree to which there is an imbalance in the true-false keying of the items in the scale (Couch & Keniston, 1960; Fricke, 1956; Messick & Jackson, 1961b; Wiggins, 1959).

Imbalance of true-false keying in a given scale creates an additional problem in interpreting the scale. Hanley (1961) pointed out that acquiescent response sets and social desirability response sets are often confounded this way. For example, if a majority of the items in a scale are keyed

"true" and simultaneously have socially desirable (or undesirable) scale values, then it is difficult to isolate the influence of the tendency to acquiesce from that of the tendency to give SD (or SUD) responses. Similar considerations apply to scales in which a majority of the items is keyed "false" and simultaneously have socially desirable (or undesirable) scale values.

The basic method for demonstrating the presence of the acquiescence response set has been to show that a given questionnaire, say the California F Scale (Adorno et al., 1950), has a disproportionate number of items keyed "agree" (or true or yes). Then the "agree" items are reversed in direction and scored for disagreement. A correlation is then computed between scores on the original and reversed items. Failure to find a high negative correlation is taken as an indication of the tendency to acquiesce.

Welsh's R Scale

The R Scale of the MMPI was developed by Welsh (1956) as a marker variable for the second MMPI factor. It consists of 40 items all of which are keyed "false." Individuals with high rates of "true" responding would obtain low scores on the R Scale, and individuals with low rates of "true" responding would obtain high scores. Consequently, scores on the R Scale have been interpreted as reflecting acquiescent tendencies. It has been suggested (Hanley,

1961) that scales containing items with neutral SDSVs might be relatively independent of the influence of SD tendencies, and therefore more susceptible to the influence of acquiescent tendencies. While the SDSVs of the items on the R Scale are not all neutral, they are distributed fairly well over the social desirability continuum, and the scale has a fair degree of balance in its SD-SUD keying (Edwards, 1964). Additionally, the R Scale tends to have a low correlation with the Edwards SD Scale, a low loading on the first MMPI factor, and a high loading on the second factor. Furthermore, the factor loadings of MMPI scales on the second factor are substantially correlated ($r = 0.82$) with the proportion of items keyed "true" in the scales, whereas the correlation between the second factor loadings and the proportion of items keyed for SD responses is relatively low ($r = -0.38$). This evidence suggests that the second MMPI factor is reflecting acquiescent tendencies (Edwards, 1970).

Infrequency

Infrequency scales are used to detect unusual or atypical ways of answering the test items. A familiar example of this type of scale is the F scale of the MMPI. Unlike most of the other MMPI scales, the F scale was not derived by comparing item endorsements between criterion and normal groups. It consists of 64 items that no more than 10 percent of an early subsample of the MMPI normative sample

answered in the deviant direction, and therefore is often referred to as the frequency (or infrequency) scale of the MMPI. The scale taps a wide variety of content areas, including bizarre sensations, strange thoughts, peculiar experiences, feelings of isolation and alienation, and a number of unlikely or contradictory beliefs and self-descriptions (Greene, 1980). Examples of F scale items with the deviant answer indicated in parentheses are as follows:

"When I am with people, I am bothered by hearing very queer things." (true)

"No one cares much what happens to you." (true)

"I believe in law enforcement." (false)

High scores on the F Scale may be obtained by answering the items without reading them carefully, (thus raising concern about the validity of the profile). They may also be obtained by persons experiencing severe psychological distress. However, since the content of the F scale items is obvious, individuals may lower or raise their F scale scores virtually as desired. Therefore, another possible interpretation of high scores on this scale is that the individuals are intentionally wishing to call attention to themselves by giving many unusual and improbable responses.

Jackson Infrequency Scale.

The 20 item Jackson Infrequency Scale, items for which are similar to items on the F scale, can be used in con-

junction with other scales to identify careless, nonpurposeful, invalid responding (Sechrest & Jackson, 1963). Writing items for the scale again involved identifying facets of behavior that are so common or so uncommon in the general population that a purposeful keyed response in the infrequent direction would be exceptionally rare. It differs from the F Scale in that items too extreme in desirability, or too bizarre or ludicrous, were eliminated. Items whose content reflected statistical deviancy, yet were neither bizarre nor particularly undesirable (i.e. "I learned to repair watches in Switzerland.") were retained. A set of 165 items showing little content homogeneity, apart from representing infrequently endorsed content, was item analyzed by correlating each item with the total score, with each of several content scales from the Personality Research Inventory (Jackson, 1976), and with a social desirability scale. Items were retained which correlated higher with the Infrequency Scale than with any of the other scales. This analysis eliminated items which reflected a predominance of desirability variance, and items that correlated inappropriately with irrelevant scales. Of the original 165 items, 40 survived this procedure, were balanced with respect to "true" and "false" responses, and assigned to parallel forms by computer, yielding two scales of 20 items length with virtually identical statistical properties (Jackson, 1976).

These infrequency items, although diverse in content, elicit reliable responses and are highly sensitive to certain types of invalid responding.

Impression Management or "Faking"

In their discussion of the K scale, Meehl and Hathaway (1946) have observed that one of the most important defects of personality scales is their susceptibility to "faking," that is, to motivation on the part of some individuals to create either a favorable or unfavorable impression of themselves. Faking refers to a conscious and intentional falsification of responses to personality items.

A person's needs may dispose him or her to present himself or herself in a particular light in a testing situation. If it is important for the individual to gain approval or acceptance, to deny inadequacies, or to achieve recognition or status, one may anticipate that his or her test responses will be altered to serve these aims. As Rotter (1960) put it, "What we call faking is only our recognition of the fact that the subject is taking the test with a different purpose or goal than the one the experimenter wants him to have" (p.308).

Faking good refers to the attempt to create a favorable impression by falsifying responses to personality items. If an individual attempts to fake good and is successful, then we may expect him or her to obtain a higher score on a scale

than he or she otherwise would, provided the trait being measured by the scale is a socially desirable one. On a scale measuring a socially undesirable trait, a successful attempt to fake good should result in a lower score than would otherwise be obtained by the individual.

Faking bad, on the other hand, refers to motivation to create an unfavorable impression. Thus the influence of faking bad on personality scales should be just the opposite of faking good. Faking bad may occur in situations where an individual is under the impression that portraying himself in terms of socially undesirable characteristics may work to his advantage. To fake bad in some situations may be of importance, but it does not seem that this form of impression management would be as prevalent as that to fake good.

The procedure frequently used in studies of faking is to have subjects take an inventory under standard instructions for self-description, and then to obtain a second record from the same subjects under special instructions to fake, that is, asking the subjects to create a favorable or unfavorable impression. General instructions to fake good on an inventory containing many scales cannot be expected to have much influence on items with extreme SDSVs. If an item has a high SDSV, then the proportion of individuals answering "true" to the item under standard instructions will also be quite high, and at best this proportion can, under the

special instructions, be increased only slightly. The same considerations apply to items with low SDSVs where the proportion of "false" responses under standard instructions is high.

For neutral items, and items with moderately desirable and undesirable scale values, the proportion of individuals answering "true" or "false" to items under standard instructions tends to be more evenly balanced. Therefore, for items of this kind there exists a greater possibility of large shifts in the proportion answering "true" or "false" under specific instructions to fake good. However, there are individual differences in the judged desirability of items. Such individual differences will not result in a uniform shift in a consistent direction on the part of all individuals responding to the items under special instructions to fake good. It seems reasonable to assume that under standard instructions subjects will differ in their tendencies to respond to items in personality inventories in terms of social desirability. If a subject's responses under standard instructions are already primarily influenced by considerations of social desirability, special instructions to attend to social desirability should not result in any marked shifts in his responses. On the other hand, subjects whose initial responses to the items under standard instructions were relatively mildly influenced by

social desirability considerations might be expected to show marked shifts in their responses under the special instructions. Faking good on personality inventories, without special instructions to do so, should be considered equivalent to the tendency to give socially desirable responses in self-description. It is this tendency that the SD scale was designed to measure.

To the degree to which individuals can accurately judge the SDSVs of the items, and to the degree to which there is an imbalance in the SD-SUD keying of the items in the various scales included in an inventory, the mean scores on the scales may be expected to change under specific instructions to fake good. For example, if all of the items in a scale in an inventory are keyed for trait and SD responses, then all individuals who accurately judge the SDSVs of the items, and in turn give the SD responses to the items, should obtain a score under special instructions that is equal to the number of items in the scale. As a result, the mean score on the scale should be increased. But if the scale contains a considerable number of items with neutral or moderately desirable or undesirable scale values, then the shift in the mean score under special instructions to fake good may not be as great as would be the case if the items were more extreme in their SDSVs. If a scale contains a balance in its SD-SUD keying, then instructions to fake

good should result in a mean score on the scale that is approximately equal to one-half the number of items in the scale.

There is no way in which to differentiate between those individuals who obtain high (or low) scores on personality scales designed to measure specific traits because of impression management, and those who obtain high (or low) scores because they have accurately described themselves. One could take the extreme position, as does Edwards, and regard scores on any scale as simply resulting from individual differences in motivation toward impression management with respect to the particular trait being measured by the scale. When one considers some of the factors involved in successful impression management, it is difficult to believe that reliable individual differences in scores on trait scales could simply result from reliable individual differences in impression management. The individual motivated toward impression management must consider the impression he wishes to create, the possible traits measured by the various scales in the inventory, the particular items included in these scales, and the keyed responses to these items. In addition, if he believes that either extremely high or low scores on a scale are to be avoided, but that a score slightly above or below average is desirable, then he must also attempt to guess the mean score on the scale. The

problems that individuals motivated toward impression management must face in responding to items in personality inventories are not inconsiderable if they are to be successful. To consider individual scores on a personality scale simply as a result of individual differences in impression management does not seem reasonable.

To summarize thus far, the major response sets adversely affecting the validity and usefulness of trait scales can be referred to as Edwards social desirability, Marlowe-Crowne social desirability, acquiescence, random responding, and impression management.

Introduction to the Chapman Scales

Researchers investigating the etiology of schizophrenia often use subjects who are considered to be at risk for the illness but who have not yet developed the disorder. If these high risk subjects could be identified more reliably than is now possible, the progress of such research would be facilitated. Because the genetic transmission of a predisposition to this disorder has been established, the relatives of those having the illness have been used as high risk subjects. However, most relatives never develop these disorders. Thus, methods for identifying high risk subjects within the general population are needed to permit the study of such persons prior to the onset of illness or, at least, much earlier in its course.

Chapman et al. (1976, 1978, 1983, 1984) have reported the construction and validation of numerous self-report inventories designed to measure personality traits thought to be associated with predisposing factors to schizophrenia. The Chapman scales that are receiving the most acceptance and application are the following: Physical Anhedonia, Perceptual Aberration, Magical Ideation, and Impulsive Nonconformity. While Chapman claims that in the development of these scales the "potential artifacts of social desirability, acquiescence, and random responding were ruled out," (1976, p. 374) close inspection of the methods of construction and item content strongly suggests that the problems of response biases were not adequately addressed in the development of the scales. A number of examples serve to illustrate this point.

First, Chapman et al., in developing this series of scales, did not use instruments designed to measure the various response biases in a consistent fashion. For instance, a different scale was used to measure acquiescence in each study, making comparisons difficult. Two of the scales used, one measuring acquiescence and the other infrequency, were, in fact, experimental scales (Chapman et al., 1976, 1983).

Second, when addressing acquiescence on both the Perceptual Aberration and Impulsive Nonconformity scales,

Chapman dismisses the problem of this response set by stating, "The number of items was not balanced because the statement of the absence of a trait could often be achieved only by inserting a "not" or "never" or some other negative creating a double negative which was likely to be confusing" (1984, p. 683).

Third, in developing the Physical Anhedonia Scale, Chapman simply remarks that, "an attempt was made to avoid items with a bias toward social desirability" (1978, p. 376) and then neglects to report a correlation with any social desirability scale.

Finally, and most importantly, while Chapman et al. (1976, 1978, 1983, 1984) address the possibility of Marlowe-Crowne social desirability response bias in their description of their scales, they neglect to deal with the problems associated with Edwards social desirability which, as we have seen, is the major source of response bias in personality measurement.

To repeat, it is necessary to have a scale as unbiased as possible with regard to SD, and certain precautions are mandatory. Many scales, bearing a number of different labels, are interpretable in terms of response biases. These scales often intercorrelate much too highly at the expense of efficiency of assessment and convergent and discriminant validity. In light of the evidence for the

massive effects of such general sources of bias as desirability and acquiescence, it is mandatory to incorporate methods of control in the development of personality scales. If a test is designed with the hope of measuring the trait of concern accurately, response biases cannot be permitted to run rampant.

If we wish to distinguish between the traits of perceptual aberration, magical ideation, impulsive nonconformity, and physical anhedonia, then it is necessary to develop scales which are relatively uncorrelated with scores on SD scales. There is no good reason why psychologists and other research workers should continue to develop new personality scales, and to assign different trait names to these scales, only to find that scores on the various scales are highly correlated with individual differences in rates of SD responding. Of much greater importance and significance would be the development of scales with sound psychometric properties on which scores are relatively uncorrelated with rates of SD responding.

The purpose of this investigation was to determine the amount of variance in the Chapman scales that could be attributed to response bias. The contention being advanced is that the Chapman scales are simply measuring, to a considerable extent, the tendency of subjects to give socially desirable responses in self-description. In this study, the

Chapman scales were re-examined for possible response biases in a systematic fashion using established instruments. Current normative data was gathered for the Edwards and Marlowe-Crowne social desirability scales, Welsh's R Scale, and Jackson's Infrequency Scale, as well as information on the ability of subjects to fake their responses on the Chapman scales.

Method

Subjects

Subjects (N = 228) were randomly selected from male and female college students who completed the Chapman scales (Perceptual Aberration, Magical Ideation, Physical Anhedonia, and Impulsive Nonconformity Scales) as part of a general screening in the 110 Introduction to Psychology course at the University of Montana. Subjects received credit for participating in a psychological experiment as a supplement to their required coursework in Psychology 110.

Procedures

Subjects first completed the Chapman scales during the regularly scheduled class period. Subjects subsequently arranged to complete the Edwards and Marlowe-Crowne social desirability scales, Welsh's R Scale, and Jackson's Infrequency Scale with the experimenter. At that time they were also randomly assigned to fake either good or bad on the Chapman scales. The faking procedure was done according to standard instructions used in studies investigating impression management. (See Appendix E).

Data analysis consisted of multiple regression and correlation techniques designed to determine the amount of variance in the Chapman scales that could be attributed to response sets, and to random responding. Additionally, measures of central tendency were calculated for the Edwards and Marlowe-Crowne social desirability scales, Welsh's R

Scale, and Jackson's Infrequency Scale, and t-tests were performed to determine if subjects were able to accurately identify the keyed responses on the Chapman scales, and therefore increase or decrease their scores on these scales as desired.

Results

Correlational Analysis

A total of 228 subjects (139 women, 89 men) participated in the study. The means and standard deviations were computed for the Edwards and Marlowe-Crowne social desirability scales, Welsh's R Scale, and Jackson's Infrequency Scale, and are presented in Table 2. The means and standard deviations obtained by the original investigators (Edwards, 1963; Marlowe & Crowne, 1964; Welsh, 1956; Jackson, 1967) are given in parentheses.

Table 2.

Current and past normative data for the Edwards (ESD) and Marlowe-Crowne social desirability (MCSD) scales, Welsh's R Scale (WR), and Jackson's Infrequency Scale (JI).

<u>Scale</u>	<u>Mean</u>	<u>St. Dev.</u>	<u>No. Items</u>
ESD: Males	26.66 (32.39)	5.53 (5.20)	39
Females	25.65	4.37	
MCSD: Males	12.55 (15.06)	4.62 (5.58)	33
Females	12.97 (16.82)	4.74 (5.50)	
WR: Males	14.72 (14.95)	4.29 (5.03)	40
Females	15.68 (15.59)	3.68 (4.51)	
JI: Males	0.29 (0.69)	0.55 (1.09)	20
Females	0.27 (0.53)	0.53 (0.79)	

Males (N = 139)

Females (N = 89)

1. Mean of combined male and female sample.

2. Prorated on the basis of partial norms.

The mean scores for both men and women on the Edwards' SD Scale were significantly lower ($t(426) = -10.99$, $p < .005$ and $t(426) = -14.56$, $p < .005$, respectively) than the values obtained by Edwards (1963). The mean scores on the Marlowe-

Crowne SD Scale were significantly lower than the values obtained by Marlowe and Crowne (1964) for both men and women ($t(892) = -9.08, p < .005$ and $t(978) = -11.77, p < .005$, respectively). The mean scores on Jackson's Infrequency Scale were also significantly lower than the original values obtained by Jackson (1967) for men and women ($t(1257) = -5.25, p < .005$ and $t(1228) = -3.44, p < .005$, respectively).

Scores for the above response set measures ranged as follows: the Edwards Social Desirability Scale ranged from 10 to 36, the Marlowe-Crowne Social Desirability Scale ranged from 1 to 27, the Welsh's R Scale ranged from 5 to 27, and Jackson's Infrequency Scale ranged from 0 to 2.

Table 3 presents the Pearson product-moment correlations among the response set measures.

Table 3.

Correlations between the Edwards and Marlowe-Crowne social desirability scales, Welsh's R Scale, and Jackson's Infrequency Scale.

	<u>Edwards SD</u>	<u>M-C SD</u>	<u>Welsh's R</u>
M-C SD	.24		
Welsh's R	.12	.18	
Jackson's Inf.	-.15	.11	-.20

(N = 220)

The means and standard deviations for scores on the Chapman scales given under standard instructions are

displayed in Table 4, and can be compared to the means and standard deviations obtained by Chapman et al. (1976, 1978, 1983, 1984) for college populations by looking at the values in parentheses. (Standard deviations for the Impulsive Nonconformity Scale were not provided by Chapman (1984) in his description of the scale). Scores on the Perceptual Abberation Scale ranged from 0 to 32 with higher scores indicating the experience of more frequent unusual perceptions. Scores on the Magical Ideation Scale ranged from 1 to 26 with higher scores suggesting a stronger belief in magical influences. Physical Anhedonia scores ranged from 1

Table 4.

Means and Standard Deviations for the Perceptual Aberration Scale (Per. Ab.), the Magical Ideation Scale (Mag. Id.), the Physical Anhedonia Scale (Phy. An.), and the Impulsive Nonconformity Scale (Imp. Non.) under Standard Instructions. Values in parentheses are from Chapman's Original reports.

<u>Scale</u>		<u>Mean</u>		<u>St. Dev.</u>	<u>No.Items</u>	
Per. Ab.:	Males	6.19	(5.97)	5.62	(5.80)	35
	Females	8.47	(7.35)	6.17	(6.60)	
	Combined	7.58		6.05		
Mag. Id.:	Males	9.84	(8.56)	5.43	(5.24)	30
	Females	12.08	(9.69)	5.57	(5.93)	
	Combined	11.21		5.61		
Phy. An.:	Males	13.46	(7.00)	6.70	(3.90)	61
	Females	9.95	(5.60)	5.62	(3.50)	
	Combined	11.32		6.29		
Imp. Non.:	Males	16.73	(16.83)	8.99		51
	Females	14.93	(12.24)	7.47		
	Combined	15.63		8.13		

Males (N = 89)

Females (N = 139)

to 37, higher scores indicating an absence of positive physical sensations. Scores ranging from 0 to 43 were obtained from responses to the Impulsive Nonconformity Scale, higher scores suggesting impulsive, defiant behavior. The above ranges were sufficiently large to assume adequate variability on these measures for the purposes of the proposed analysis.

The means obtained for the Magical Ideation Scale for both men and women were significantly higher ($t(771) = -4.95, p < .005$ and $t(967) = -8.83, p < .005$, respectively) than those reported by Chapman (1984) and his colleagues. The means for the Physical Anhedonia Scale for both men and women were also significantly higher ($t(429) = -27.4, p < .005$ and $t(402) = -20.6, p < .005$, respectively) than those reported by Chapman et al.

There were basically six observations that were anomalies with respect to the entire set of data. Some were specific to a single scale because of an extreme score on the scale, while others were outliers with respect to two or more of the measures. In each case, the corresponding extreme scores were deleted for computing correlations with scores from the Chapman scales, and for the multiple regression analysis for each of the Chapman scales (presented below) to obtain a more realistic impression of the findings (Snedecor & Cochran, 1983). Thus, the correlations in Table

5 between each of the Chapman scales are presented as average correlations. The reader is referred to Appendix G for the actual correlations resulting from omitting observations from each set of data. Table 6 gives the correlations between the response scale measures and each of the Chapman scales.

Table 5.

Average Correlations between the Perceptual Abberation Scale (Per. Ab.), the Magical Ideation Scale (Mag. Id.), the Physical Anhedonia Scale (Phy. An.), and the Impulsive Nonconformity Scale (Imp. Non.).

	<u>Per. Ab.</u>	<u>Mag. Id.</u>	<u>Phy. An.</u>
Mag. Id.	.683		
Phy. An.	-.171	-.173	
Imp. Non.	.354	.353	.226

(N approximately 225)

Table 6.

Correlations between the Chapman Scales and the Response Set Measures.

	<u>Edward's SD</u>	<u>M-C SD</u>	<u>Welsh's R</u>	<u>Jackson's Inf.</u>
Per. Ab.	-.359	-.067	-.193	.279
Mag. Id.	-.428	-.081	-.292	.242
Phy. An.	-.118	-.083	.038	.110
Imp. Non.	-.382	-.098	-.300	.195

(N = 228)

Multiple Regression Analysis

To further assess the influence of response sets on the

Chapman scales, a stepwise multiple regression analysis was performed for each of the four Chapman scales using scores on the four response set measures as predictor variables. Values of F to enter and remove a variable were 4.0 and 2.0, respectively. These correspond approximately to the .01 and .05 levels of significance. For the regression analysis, subjects were divided into two groups: an original sample consisting of 150 subjects (90 females and 60 males), and a cross validation sample consisting of 78 subjects (49 females and 29 males). The analysis of the original sample will be described first.

Table 7 presents the results of the stepwise regression analyses describing those variables that contributed significantly to the prediction of each of the Chapman scales. For the Perceptual Abberation Scale, the two strongest predictors were the Edwards Social Desirability Scale and Jackson's Infrequency Scale, which together accounted for 17.6 percent of the total variance ($F(2,142) = 8.76, p < .0005$).

The strongest predictors for the Magical Ideation Scale were the Edwards Social Desirability Scale and Welsh's R Scale, which in combination accounted for 25.1 percent of the total variance ($F(2,140) = 13.04, p < .0005$).

For the Physical Anhedonia, no reliable predictors were identified by the stepwise procedure. The four response set

Table 7.

Prediction of Chapman Scales from Response Set Measures-
Stepwise Regression Analysis.

<u>Scale</u>	<u>Predictors</u>	<u>% of Variance/ each predictor</u>	<u>Regr. Weight</u>	<u>Cond.t*</u>
Per. Ab.	ESD	17.6/12.9	-.37	-4.13
	JI	and 4.7	2.58	2.86
Mag. Id.	ESD	25.1/18.4	-.47	-5.61
	WR	and 6.7	-.39	-3.59
Imp. Non.	ESD	22.4/14.6	-.58	-4.98
	WR	and 7.8	-.57	-3.80
Phy. An.	JI	2.1/ 1.2	1.79	NS
	ESD	-	-.06	NS
	MCSD	-	-.07	NS
	WR	-	.13	NS

* df = approximately 220, p < .01

measures together accounted for only 2.1 percent of the total variance ($F(4,141) = .76$, p > .05), indicating that individuals were not responding to items on this scale in terms of response sets.

For the Impulsive Nonconformity Scale, the two strongest predictors were again Edwards Social Desirability Scale and Welsh's R Scale which together accounted for 22.4 percent of the total variance ($F(2,141) = 10.55$, p < .0005).

Cross Validation

Cross validation of the regression equations for each of the Chapman scales resulted in considerable shrinkage in the amount of variance accounted for by the identified

response set predictor variables. For the Perceptual Abberation Scale, the percentage of predictable variance decreased from 17.6 percent to 2.8 percent; for the Magical Ideation Scale, from 25.1 percent to 12.2 percent; and for the Impulsive Nonconformity Scale, from 22.4 percent to 9.3 percent. This amount of shrinkage was much greater than would generally be expected from the cross validation procedure.

Consequently, Norman's double-split cross validation procedure was used to further assess whether the sample was atypical, or whether the relationships identified by the regression equations were unstable. Subjects were divided into two groups, Sample A and Sample B. For both samples ($N = 114$), stepwise regressions were performed for each of the Chapman scales on the four response set measures, and the alternate sample served as the cross validation sample. As can be seen from the data in Table 8, the procedure resulted in the selection of different variables as predictors for the two samples, and again the shrinkage was larger than expected, in two cases alarmingly great.

Analysis of Faking

To investigate the ability of subjects to fake their responses on the Chapman scales, all subjects were randomly assigned to either fake good or to fake bad on the Perceptual Abberation, Magical Ideation, Physical Anhedonia, and

Impulsive Nonconformity Scales. The number of subjects instructed to fake good was 115, the number who faked bad was 113. The means and standard deviations for scores faked in both directions are presented in Table 9.

In comparing the means from the faking and standard instructions we find that, on the Perceptual Abberation

Table 8.

Double-Split Cross Validation of Stepwise Regression of the Chapman Scales on the Response Set Measures

<u>Sample A</u>	<u>Identified Predictors</u>	<u>% of Variance in Sample A</u>	<u>% of Variance in Sample B</u>
Per. Ab.	ESD, JI, WR	23.4	3.3
Mag. Id.	ESD, WR	21.6	15.7
Imp. Non.	ESD, WR	23.0	13.5
<u>Sample B</u>	<u>Identified Predictors</u>	<u>% of Variance in Sample B</u>	<u>% of Variance in Sample A</u>
Per. Ab.	ESD	18.9	9.1
Mag. Id.	ESD	18.3	9.8
Imp. Non.	MCSD, WR, ESD	21.3	2.6

df = approximately 220, $p < .01$

Scale, subjects in the fake good condition were able to lower the average score 1.8 points (0.31 standard deviations) from their scores obtained under standard instructions ($t(114) = 2.95$, $p < .01$). Subjects in the fake bad condition were able to raise the average score 18.3 points (3.02 standard deviations) from their scores obtained under standard instructions ($t(112) = -15.92$, $p <$

.0001). The asymmetry in the amounts lowered and raised, respectively, are of course related to the skewness of the original distribution.

On the Magical Ideation Scale, subjects in the fake good condition were able to lower the average score 5.2

Table 9.

Means and Standard Deviations for the Chapman Scales under Faking Instructions.

<u>Scale</u>	<u>Fake Good</u>		<u>Fake Bad</u>	
	<u>Mean</u>	<u>St. Dev.</u>	<u>Mean</u>	<u>St. Dev.</u>
Per. Ab.	5.77	7.38	25.87	10.27
Mag. Id.	6.00	5.51	24.61	6.83
Phy. An.	8.72	10.36	48.23	18.91
Imp. Non.	12.23	10.18	43.11	11.62

Fake good (N = 115)

Fake bad (N = 113)

points (0.93 standard deviations) from their scores obtained under standard instructions ($t(114) = 10.51, p < .0001$). Subjects in the fake bad condition were able to raise the average score 13.4 points (2.39 standard deviations) from their scores obtained under standard instructions ($t(113) = -15.99, p < .0001$). As above, asymmetry accounts for the differential shift in scores.

On the Physical Anhedonia Scale, subjects in the fake good condition were able to lower the average score 2.6 points (0.41 standard deviations) from their scores obtained

under standard instructions ($t(114) = 2.44, p < .01$). Subjects in the fake bad condition were able to raise the average score 36.9 points (5.87 standard deviations) from their scores obtained under standard instructions ($t(113) = -19.51, p < .0001$). As above, the skewness of the original distribution accounts for the asymmetry.

On the Impulsive Nonconformity Scale, subjects in the fake good condition were able to lower the average score 3.4 points (0.42 standard deviations) from their scores obtained under standard instructions ($t(114) = 4.69, p < .0001$). Finally, subjects in the fake bad condition were able to raise the average score 27.5 points (3.38 standard deviations) from their scores obtained under standard instructions ($t(113) = -19.17, p < .0001$). Again, asymmetry account for the differential shifts.

When instructed to fake good, the ranges of scores on the Chapman scales were quite wide. Scores on the Perceptual Abberation Scale ranged from 0 to 35. Scores on the Magical Ideation Scale ranged from 0 to 30. Physical Anhedonia scores ranged from 0 to 60. Scores ranging from 0 to 50 were obtained on the Impulsive Nonconformity Scale.

Under fake bad instructions, scores on the Perceptual Abberation, Magical Ideation, Physical Anhedonia, and Impulsive Nonconformity Scales ranged from 0 to 35, 3 to 30, 1 to 61, and 7 to 51, respectively. Note that the score

ranges for both fake good and fake bad conditions indicate that some individuals obtained a maximum or minimum score in the wrong direction, that is, completely reversed the faking task. Even when including such subjects in the analysis, the shifts in the mean scores under faking instructions were always significant.

We were interested to see if those individuals identified as being psychosis prone, using Chapman's criteria, were able to effectively fake their responses on the Chapman scales. Persons scoring more than two standard deviations above the mean on either the Perceptual Abberation Scale or the Magical Ideation Scale, or scoring greater than three standard deviations above the mean on the combined Perceptual Aberration - Magical Ideation Scale, are identified by Chapman (1983) et al. as "Per-Mags," (i.e. psychosis prone) Using this criterion, 14 subjects from our sample were identified as Per-Mags.

The number of Per-Mags instructed to fake good was 8, the number who faked bad was 6. The means and standard deviations for scores faked in both directions by these individuals are presented in Table 10. In comparing the means from the faking and standard instructions we find that, when faking good on the Perceptual Abberation and Magical Ideation Scales, Per-Mags were able to lower their average score 2.07 and 2.34 standard deviations, respec-

Table 10.

Means and Standard Deviations for Per-Mags on the Chapman Scales under Faking Instructions.

<u>Scale</u>	<u>Fake Good</u>		<u>Fake Bad</u>	
	<u>Mean</u>	<u>St. Dev.</u>	<u>Mean</u>	<u>St. Dev.</u>
Per. Ab.	8.25*	7.13	28.50	6.06
Mag. Id.	12.88*	7.41	25.17	4.75
Phy. An.	7.00	3.74	49.33**	21.97
Imp. Non.	16.12	8.46	43.50*	9.72

Fake good (N = 8), fake bad (N = 6)

* $p < .05$

** $p < .01$

tively. When faking bad on the Physical Anhedonia and Impulsive Nonconformity Scales, Per-Mags were able to raise their average score by 6.71 and 2.62 standard deviations, respectively. The probabilities for t-tests of these mean differences are indicated in Table 10 by means of asterisks.

The correlations for the Per-Mag group between the Chapman scales and the response set measures are presented in Table 11. The correlations between the two sets of measures for the full sample are repeated for reference. Given the very small sample of Per-Mags, the differences in the two sets of values are difficult to interpret. That is, both the more unusual behavior to be expected of Per-Mags, and sampling variation due to a small n are plausible, but competing, explanations.

Table 11.

Correlations between the Chapman Scales and the Response Set Measures for Per-Mags.

	<u>Edward's SD</u>	<u>M-C SD</u>	<u>Welsh's R</u>	<u>Jackson's Inf.</u>
Per. Ab.	-.08	-.49	.42	-.28
Mag. Id.	-.03	-.31	-.12	.32
Phy. An.	-.28	-.16	-.36	.20
Imp. Non.	.02	-.03	-.28	.43

(N = 14)

Correlations between the Chapman Scales and the Response Set Measures for entire sample.

<u>Inf.</u>	<u>Edward's SD</u>	<u>M-C SD</u>	<u>Welsh's R</u>	<u>Jackson's</u>
Per. Ab.	-.359	-.067	-.193	.279
Mag. Id.	-.428	-.081	-.292	.242
Phy. An.	-.118	-.083	.038	.110
Imp. Non.	-.382	-.098	-.300	.195

(N = 228)

Discussion

The first major finding of the study was that the means for the Edwards' and Marlowe-Crowne social desirability scales were significantly lower than the means reported in the 1960's. As the present sample was relatively large ($N = 228$), it is unlikely that this finding is the result of a sampling error or selection bias. Rather, it appears that social desirability norms are less stringent than they once were.

In addition, the mean score for Jackson's Infrequency Scale was significantly lower than originally reported in 1967. This may suggest a trend toward more sophisticated approaches to psychological tests by college populations. It may also reflect the fact that much of the present study was done in small groups with greater control over random or inattentive responding.

The normative data obtained for the Perceptual Aberration and Impulsive Nonconformity Scales was consistent with the norms obtained by Chapman and his colleagues. On the Magical Ideation Scale, however, the mean from the present Montana sample was slightly, though significantly higher than the Wisconsin sample mean obtained by Chapman and Eckblad (1983). The mean for the Physical Anhedonia Scale is also significantly higher than Chapman's value, and may be large enough to be of practical significance (6 point

difference for men, 4 point difference for women). This suggests the need for further collection of normative data reflecting possible geographical variation in responding on these scales.

The response set measures identified by the stepwise regression analysis as the strongest predictors for the Magical Ideation and Impulsive Nonconformity Scales were the Edwards Social Desirability Scale and Welsh's R Scale. These results were highly significant, accounting for approximately 25 percent of the total variance in these scales, and indicate that these two response sets are influencing scores on the Magical Ideation and Impulsive Nonconformity Scales in a substantial way.

The Edwards is again the strongest predictor for the Perceptual Aberration Scale. The Edwards SD Scale and Jackson's Infrequency Scale together accounted for 17.6 percent of the variance.

Interestingly, none of the response set measures was found to be a reliable predictor of scores on the Physical Anhedonia Scale, and correlations between the Physical Anhedonia Scale and the response set measures were uniformly low (Table 6). One can begin to see the reason for this by inspecting items in the scale. For example, there is a balance in the true-false keying of the scale (31 items keyed 'true', 30 items keyed 'false') a characteristic which

has been recommended to reduce the unwanted influence of acquiescent response sets. Secondly, it is not unreasonable to consider that admitting one cares little for a variety of physical stimuli is less socially undesirable than admitting to the experience of unusual perceptual experiences, believing in magic, or to thinking oneself to be an impulsive and defiant person.

For example, one item on the Physical Anhedonia Scale is "Sex is the most intensely enjoyable thing in life." The answer to this item clearly depends on the subject's personal values, and we can easily imagine an individual responding yes or no, and therefore we can reasonably expect variability in responses. Thus, societal agreement regarding the social desirability of signs and symptoms of physical anhedonia is probably low, and we would look for little influence from tendencies to respond in the socially desirable direction.

Another important finding in this research was that the set of predictor variables selected from the stepwise regression procedure was strongly influenced by the sample. Correspondingly, cross validation of the prediction equations often resulted in a tremendous amount of shrinkage indicating these relationships are unstable and highly dependent on the sample. Given that the response set measures used are very reliable scales, it is possible that

validation experiments for the Chapman scales have been consistently capitalizing on chance, as cross validation procedures are seldom reported. Thus, a few individuals in any sample indicating bizarre symptoms generally associated with schizotypal personality disorder may account for the magnitude of the correlations they report, and also for the lack of replication of his results.

The analysis of the faking conditions demonstrates that normal subjects were able to successfully fake their responses in either direction on each of the Chapman scales. The means for the Chapman measures are low under standard instructions, a fact which by itself suggests that social desirability for these scales is not adequately controlled. While the ranges of the scores on the Chapman scales under faking instructions indicated that occasional individuals reversed the faking task, normal subjects were, in general, able to see the structure of the scales clearly enough to be able to create a good or bad impression as desired for each of the Chapman measures.

Using Chapman's criteria, 14 Per-Mags were identified in our sample. Note that such a small number is to be expected given the selection criteria for Per-Mags. Per-mags in our sample were able to significantly lower their score when faking good on the Perceptual Aberration and Magical Ideation Scales, and significantly raise their

scores when faking bad on the Physical Anhedonia and Impulsive Nonconformity Scales. They were not able to significantly alter their scores in other situations. Thus, it appears that the Per-Mag group was not able to alter their impressions to the same extent as the normal group. However, it is difficult to determine if this was a result of initially low mean values (even when they altered their scores successfully the means were still greater than the average values for the scales), of the small sample (in which case the event of an individual reversing the faking task would produce tremendous shifts in the mean values), or whether these individuals were themselves unable to fake effectively.

In conclusion, it is extremely important to develop personality scales that are relatively free from response biases to ensure that the scales are not simply measuring personal sets to respond to content-irrelevant aspects of the test. While Chapman et al. claim that the potential artifacts of social desirability, acquiescence, and random responding were considered, and thereby accounted for, in their psychosis-proneness measures, the inconsistent manner with which response biases were measured made evaluation of this claim quite difficult. Closer, more systematic inspection reveals that the problems of response biases were not fully addressed.

Chapman et al. address the possibility of Marlowe-Crowne social desirability response bias in the descriptions of their scales. However, they neglect to deal with the problems associated with Edwards social desirability which is the major source of response bias in personality measurement. This is reflected in the fact that the Edwards SD Scale was found to be the strongest predictor for the Perceptual Aberration, Magical Ideation, and Impulsive Nonconformity Scales, and the fact that normal subjects were able to fake their responses either direction on each of the Chapman measures.

If we wish to distinguish between the traits of perceptual aberration, magical ideation, impulsive nonconformity, and physical anhedonia, then it is necessary to develop scales which are relatively uncorrelated with scores on both Edwards and Marlowe-Crowne social desirability scales. In addition, since scores on personality scales are regarded as being susceptible to the influence of acquiescent tendencies to the degree to which there is an imbalance in the true-false keying of the items, an attempt to balance the keying of the Chapman scales, (a relatively simple task), would result in a diminished influence of acquiescence response sets. These problems must be adequately addressed before reasonable interpretations of the Chapman scales can be made. Additionally, the faking analysis should be

replicated with a larger Per-Mag sample to determine the extent to which individuals identified as psychosis prone can fake their responses on the Chapman measures.

References

- Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). The authoritarian personality. New York: Harper.
- Bass, B. M. (1955). Authoritarianism or acquiescence? Journal of Abnormal and Social Psychology, 51, 616-623.
- Block, J. (1965). The challenge of response sets. New York: Appleton-Century-Crofts.
- Chapman, L. J., & Campbell, D. T. (1957). Response set in the F scale. Journal of Abnormal and Social Psychology, 54, 129-132.
- Chapman, L. J., & Eckblad, M. (1983). Magical ideation as an indicator of schizotypy. Journal of Consulting and Clinical Psychology, 51(2), 215-225.
- Chapman, L. J., Chapman, J. P., Numbers, J. S., Edell, W. S., Carpenter, B. N., & Beckfield, D. (1984). Impulsive nonconformity as a trait contributing to the prediction of psychotic-like and schizotypal symptoms. Journal of Nervous and Mental Disease, 172(11), 681-691.
- Chapman, L. J., Chapman, J. P., & Raulin, M. L. (1976). Scales for physical and social anhedonia. Journal of Abnormal Psychology, 85(4), 374-382.
- Chapman, L. J., Chapman, J. P., & Raulin, M. L. (1978). Body-image aberration in schizophrenia. Journal of Abnormal Psychology, 87(4), 399-407.
- Couch, A., & Keniston, K. (1961). Yeasayers and naysayers: Agreeing response set as a personality variable. Journal of Abnormal and Social Psychology, 60, 151-174.
- Cowen, E. L., & Tongas, P. (1959). The social desirability of trait descriptive terms: Applications to a self-concept inventory. Journal of Consulting Psychology, 23, 361-365.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281-302.
- Crowne, D. P., & Marlowe, D. (1964). The approval motive: Studies in evaluative dependence. New York: John Wiley & Sons.

- Cruse, D. B. (1966). Social desirability responses at ages 3 through 6. Child Development, 37, 909-916.
- Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. Journal of Applied Psychology, 37, 90-93.
- Edwards, A. L. (1957). The social desirability variable in personality assessment and research. New York: Dryden Press.
- Edwards, A. L. (1961). Social desirability or acquiescence in the MMPI? A case study with the SD scale. Journal of Abnormal and Social Psychology, 63, 351-359.
- Edwards, A. L. (1963). A factor analysis of experimental social desirability and response set scales. Journal of Applied Psychology, 47, 308-316.
- Edwards, A. L. (1964). The assessment of human motives by means of personality scales. In D. Levin (Ed.), Nebraska symposium on motivation. Lincoln: University of Nebraska Press, 135-162.
- Edwards, A. L. (1970). The measurement of personality traits by scales and inventories. New York: Holt, Rinehart, and Winston.
- Edwards, A. L., Diers, C. J., & Walker, J. N. (1962). Response sets and factor loading on sixty-one personality scales. Journal of Applied Psychology, 46, 220-225.
- Edwards, A. L., & Walsh, J. A. (1963a). Relationships between various psychometric properties of personality items. Educational and Psychological Measurement, 23, 227-238.
- Edwards, A. L., & Walsh, J. A. (1963b). The relationship between the intensity of the social desirability keying of a scale and the correlation of the scale with the Edwards' SD scale and the first factor loading of the scale. Journal of Clinical Psychology, 19, 200-203.
- Edwards, A. L., & Walsh, J. A. (1964). Response sets in standard and experimental personality scales. American Educational Research Journal, 1, 52-61.
- Fricke, B. G. (1956). Response set as a suppressor variable

- in the OAIS and MMPI. Journal of Consulting Psychology, 20, 161-169.
- Fordyce, W. E. (1956). Social desirability in the MMPI. Journal of Consulting Psychology, 20, 171-175.
- Greene, R. L. (1980). The MMPI: An interpretive manual. Orlando, Florida: Grune & Stratton, Inc.
- Hanley, C. (1956). Social desirability and responses to items from three MMPI scales: D, Sc, and K. Journal of Applied Psychology, 40, 324-328.
- Hanley, C. (1961). Social desirability and response bias in the MMPI. Journal of Consulting Psychology, 25, 13-20.
- Jackson, D. N. (1960). Stylistic response determinants in the California Psychological Inventory. Educational and Psychological Measurement, 10, 339-346.
- Jackson, D. N. (1967). Personality research form manual. Goshen, New York: Research Psychologists Press.
- Jackson, D. N. (1976). Jackson personality inventory manual. Port Huron, Mich.: Research Psychologists Press.
- Jackson, D. N., & Messick, S. J. (1958). Content and style in personality assessment. Psychological Bulletin, 55, 243-252.
- Jackson, D. N., & Messick, S. J. (1961). Acquiescence and desirability as response determinants on the MMPI. Educational and Psychological Measurement, 21, 771-792.
- Jackson, D. N., & Messick, S. J. (1962a). Response styles and the assessment of psychopathology. In S. J. Messick & J. Ross (Eds.), Measurement in Personality and Cognition. New York: Wiley.
- Jackson, D. N., & Messick, S. J. (1962b). Response styles on the MMPI: Comparison of clinical and normal samples. Journal of Abnormal and Social Psychology, 65, 285-299.
- Kenny, D. T. (1956). The influence of social desirability on discrepancy measures between real self and ideal self. Journal of Consulting Psychology, 20, 318-318.
- Meehl, P. E., & Hathaway, S. R. (1946). The K factor as a suppressor variable in the Minnesota Multiphasic

- Personality Inventory. Journal of Applied Psychology, 30, 525-564.
- Merrill, R. M., & Heathers, L. B. (1956). The relation of the MMPI to the Edwards Personal Preference Schedule on a college counseling center sample. Journal of Consulting Psychology, 20, 310-314.
- Messick, S., & Jackson, D. N. (1961a). Desirability scale values and dispersions for MMPI items. Psychological Reports, 8, 409-414.
- Messick, S. & Jackson, D. N. (1961b). Acquiescence and the factorial interpretation of the MMPI. Psychological Bulletin, 58, 299-304.
- Rotter, J. B. (1960). Some implications of a social learning theory for the prediction of goal directed behavior from testing procedures. Psychological Review, 67, 301-316.
- Sechrest, L. B., & Jackson, D. N. (1963). Deviant response tendencies: Their measurement and interpretation. Educational and Psychological Measurement, 23, 33-53.
- Snedecor & Cochran (1983). Statistical Methods for Research Workers (8th ed.). Ames, Iowa: Iowa State University Press.
- Taylor, J. B. (1959). Social desirability and MMPI performance: The individual case. Journal of Consulting Psychology, 23, 514-517.
- Walsh J. A., Tomlinson-Keasey, C., & Klieger, D. M. (1974). Acquisition of the social desirability response. Genetic Psychology Monographs, 89, 241-272.
- Welsh, G. S. (1956). Factor dimensions A and R. In G. S. Welsh & W. G. Dahlstrom (Eds.), Basic readings on the MMPI in psychology and medicine. Minneapolis: University of Minnesota Press.
- Wiggins, J. S. (1959). Interrelationships among MMPI measures of dissimulation under standard and social desirability instructions. Journal of Consulting Psychology, 23, 419-427.
- Wiggins, J. S. (1962). Strategic, method, and stylistic variance in the MMPI. Psychological Bulletin, 59, 224-242.

Appendix A

Edwards Social Desirability Scale

Read each statement and decide whether it is "true" as applied to you, or "false" as applied to you. Indicate your answer by circling the T or F. Be sure to answer every question.

1. My hands and feet are usually warm enough. (T)
2. I find it hard to keep my mind on a task or job. (F)
3. Most of the time I would rather sit and daydream than do anything else. (F)
4. My sleep is fitful and disturbed. (F)
5. My family does not like the work I have chosen (or the work I intend to choose for my life work). (F)
6. I am happy most of the time. (T)
7. I am very seldom troubled by constipation. (T)
8. I am liked by most people who know me. (T)
9. I cry easily. (F)
10. I do not tire quickly. (T)
11. I frequently notice my hand shakes when I try to do something. (F)
12. Criticism or scolding hurts me terribly. (F)
13. It makes me impatient to have people ask my advice or otherwise interrupt me when I am working on something important. (F)
14. I dream frequently about things that are best kept to myself. (F)
15. I sweat very easily even on cool days. (F)
16. I have had periods in which I carried on activities without knowing later what I had been doing. (F)
17. It makes me uncomfortable to put on a stunt at a party even when others are doing the same sort of thing. (F)
18. I am not afraid to handle money. (T)
19. Life is a strain for me much of the time. (F)
20. I am easily embarrassed. (F)
21. I cannot keep my mind on one thing. (F)
22. When in a group of people I have trouble thinking of the right thing to talk about. (F)
23. I feel anxiety about someone or something almost all of the time. (F)
24. I have been afraid of things or people that I knew could not hurt me. (F)
25. I am not usually self-conscious. (T)
26. It does not bother me particularly to see animals suffer. (F)
27. My parents and family find more fault with me than they should. (F)
28. I feel hungry almost all the time. (F)
29. I worry quite a bit about possible misfortunes. (F)

- 30. No one cares much what happens to you. (F)
- 31. It makes me nervous to have to wait. (F)
- 32. I usually expect to succeed in the things I do. (T)
- 33. I can easily make other people afraid of me, and sometimes do so for the fun of it. (F)
- 34. I blush no more often than others. (T)
- 35. I am never happier than when alone. (F)
- 36. I shrink from facing a crisis or difficulty. (F)
- 37. I sometimes feel that I'm about to go to pieces. (F)
- 38. I have reason for feeling jealous of one or more members of my family. (F)
- 39. People often disappoint me. (F)

Appendix B

Marlowe-Crowne Social Desirability Scale

Listed below are a number of statements concerning personal attitudes and traits. Read each item and decide whether the statement is true or false as it pertains to you personally.

1. Before voting I thoroughly investigate the qualifications of all the candidates. (T)
2. I never hesitate to go out of my way to help someone in trouble. (T)
3. It is sometimes hard for me to go on with my work if I am not encouraged. (F)
4. I have never intensely disliked anyone. (T)
5. On occasion I have had doubts about my ability to succeed in life. (F)
6. I sometimes feel resentful when I don't get my way. (F)
7. I am always careful about my manner of dress. (T)
8. My table manners at home are as good as when I eat out in a restaurant. (T)
9. If I could get into a movie without paying and be sure I was not seen, I would probably do it. (F)
10. On a few occasions, I have given up doing something because I thought too little of my ability. (F)
11. I like to gossip at times. (F)
12. There have been times when I felt like rebelling against people in authority even though I knew they were right. (F)
13. No matter who I'm talking to, I'm always a good listener. (T)
14. I can remember playing sick to get out of something. (F)
15. There have been occasions when I took advantage of someone. (F)
16. I'm always willing to admit when I've made a mistake. (T)
17. I always try to practice what I preach. (T)
18. I don't find it particularly difficult to get along with loud mouthed, obnoxious people. (T)
19. I sometimes try to get even, rather than forgive and forget. (F)
20. When I don't know something I don't at all mind admitting it. (T)
21. I am always courteous, even to people who are disagreeable. (T)
22. At times I have really insisted on having things my own way. (F)
23. There have been occasions when I felt like smashing things. (F)
24. I would never think of letting someone else be punished

- for my wrongdoings. (T)
25. I never resent being asked to return a favor. (T)
26. I have never been irked when people expressed ideas very different from my own. (T)
27. I never make a long trip without checking the safety of my car. (T)
28. There have been times when I was quite jealous of the good fortune of others. (F)
29. I have almost never felt the urge to tell someone off. (F)
30. I am sometimes irritated by people who ask favors of me. (F)
31. I have never felt that I was punished without cause. (T)
32. I sometimes think when people have a misfortune they only got what they deserved. (F)
33. I have never deliberately said something that hurt someone's feelings. (T)

Appendix C

Jackson's Infrequency Scale

1. My musical compositions have been played in concert halls around the world. (T)
2. I have sight in only one eye. (T)
3. I have no sense of taste at all. (T)
4. I have kept a pet monkey for years. (T)
5. I have won trophies in professional golf tournaments. (T)
6. I run five miles every day to keep healthy (T)
7. I eat imported cheeses with all my meals. (T)
8. I have made several trips overseas to study old ruins and rock formations.
9. Everyone in my family has the same birthday. (T)
10. All jokes seem pointless to me. (T)

11. Of the people I know, I like some better than others. (F)
12. I have had at least one cold in my life. (F)
13. I have sometimes hesitated before making a decision. (F)
14. In my lifetime, I have eaten at least one meal in a restaurant. (F)
15. Some things don't turn out exactly as I plan them. (F)
16. I can eat most foods without feeling ill. (F)
17. I do somethings better than others. (F)
18. I believe there are some jobs I would not enjoy doing. (F)
19. I can walk a few blocks without getting tired. (F)
20. I usually sleep at least four hours every night. (F)

Appendix D

Welsh's R Scale

1. I like mechanics magazines. (F)
2. I like to read newspaper articles on crime. (F)
3. I am about as able to work as I ever was. (F)
4. I enjoy detective or mystery stories. (F)
5. At times I feel like smashing things. (F)
6. I am in just as good physical health as most of my friends. (F)
7. I think I would like the kind of work a forest ranger does. (F)
8. I frequently find it necessary to stand up for what I think is right. (F)
9. I like dramatics. (F)
10. I do not worry about catching diseases. (F)
11. I like to cook. (F)
12. At times I feel like picking a fist fight with someone. (F)
13. I have never had a fit or convulsion. (F)
14. I have had periods in which I carried on activities without knowing later what I had been doing. (F)
15. Sometimes, when embarrassed, I break out in a sweat which annoys me greatly. (F)
16. I like to flirt. (F)
17. I think I would like the work of a building contractor. (F)
18. I like science. (F)
19. I do not blame a person for taking advantage of someone who lays himself open to it. (F)
20. At times I am all full of energy. (F)
21. I do not often notice my ears ringing or buzzing. (F)
22. Once in a while I feel hate towards members of my family whom I usually love. (F)
23. My mother or father often made me obey even when I thought it was unreasonable. (F)
24. I have often met people who were supposed to be experts who were no better than I. (F)
25. If given the chance I would make a good leader of people. (F)
26. I like to attend lectures on serious subjects. (F)
27. I try to remember good stories to pass them on to other people. (F)
28. I was fond of excitement when I was young (or in childhood). (F)
29. I am often inclined to go out of my way to win a point with someone who has opposed me. (F)
30. I enjoy social gatherings just to be with people. (F)
31. I enjoy the excitement of a crowd. (F)
32. My worries seem to disappear when I get into a crowd of

- lively friends. (F)
33. I have had no difficulty starting or holding my urine.
(F)
34. I am often sorry because I am so cross and grouchy. (F)
35. I am fascinated by fire. (F)
36. I like to let people know where I stand on things. (F)
37. Some of my family have quick tempers. (F)
38. I would like to wear expensive clothes. (F)
39. I like repairing a door latch. (F)
40. I am very careful about my manner of dress. (F)

Appendix E

Faking Instructions

The following items are from a scale designed to measure magical ideation. This includes believing in a number of magical influences such as the transfer of psychical energies between people, psychokinetic effects, precognition, astrology, reincarnation, and good luck charms. Please try to answer these items in such a way as to give the impression that magical ideation (does not) affect(s) your thinking.

The following items are from a scale designed to measure impulsive nonconformity. This includes lacking self-control, difficulty delaying gratification, episodes of uncontrolled rage, lacking concern for the rights of others, and lacking respect for the social and ethical standards of society. Please try to answer these items in such a way as to give the impression that impulsive nonconformity (does not) affect(s) your thinking.

The following items are from a scale designed to measure perceptual aberration. This includes perceiving alterations in the size and shape of one's body, and feeling as though one is merging with external objects. Please try to answer these items in such a way as to give the impression that perceptual aberration (does not) affect(s) your thinking.

The following items are from a scale designed to measure physical anhedonia. This includes a lowered ability to experience physical pleasures such as eating, touching, movement, smelling and hearing. Please try to answer these items in such a way as to give the impression that physical anhedonia (does not) affect(s) your thinking.

Appendix F

The Chapman Scales

Perceptual Aberration Scale

3. Sometimes I have had feeling that I am united with an object near me. (T)
7. Sometimes when I look at things like tables and chairs, they seem strange. (T)
9. Sometimes I feel like everything around me is tilting. (T)
17. I can remember when it seemed as though one of my limbs took on an unusual shape. (T)
26. I have felt that my body and another person's body were one and the same. (T)
32. Parts of my body occasionally seem dead or unreal. (T)
35. Sometimes people whom I know well begin to look like strangers. (T)
40. I have sometimes had the feeling that one of my arms or legs was disconnected from the rest of my body. (T)
46. It has seemed at times as if my body was melting into my surroundings. (T)
54. I have felt as though my head or limbs were somehow not my own. (T)
58. Now and then when I look in the mirror, my face seems quite different than usual. (T)
61. Often I have a day when indoor lights seem so bright that they bother my eyes. (T)
64. I have sometimes had the feeling that my body is decaying inside. (T)
67. I have sometimes felt that some part of my body no longer belonged to me. (T)
80. I sometimes have to touch myself to make sure I'm still there. (T)
83. I have sometimes felt confused as to whether my body was really my own. (T)
88. I have felt that something outside my body was a part of my body. (T)
92. Sometimes I have had the feeling that a part of my body is larger than it usually is. (T)
95. For several days at a time I have had such a heightened awareness of sights and sounds that I cannot shut them out. (T)
96. At times I have wondered if my body was really my own. (T)
99. I sometimes have had the feeling that some parts of my body are not attached to the same person. (T)
102. My hearing is sometimes so sensitive that ordinary sounds become uncomfortable. (T)

- 107. Occasionally I have felt as though my body did not exist. (T)
- 111. Occasionally it has seemed as if my body had taken on the appearance of another person's body. (T)
- 120. I have had the momentary feeling that my body has become misshapen. (T)
- 124. Sometimes I have had a passing thought that some part of my body was rotting away. (T)
- 128. I sometimes have had the feeling that my body is abnormal. (T)
- 130. I have had the momentary feeling that the things I touch remain attached to my body. (T)
- 132. Ordinary colors sometimes seem much too bright to me (without taking drugs). (T)
- 133. Sometimes part of my body has seemed smaller than it usually is. (T)
- 136. Sometimes I have felt that I could not distinguish my body from other objects around me. (T)
- 42. My hands or feet have never seemed far away. (F)
- 51. I have never had the passing feeling that my arms or legs had become longer than usual. (F)
- 85. The boundaries of my body always seem clear. (F)
- 104. I have never felt that my arms or legs have momentarily grown in size. (F)

Magical Ideation Scale

- 5. I sometimes have a feeling of gaining or losing energy when certain people look at me or touch me. (T)
- 14. I have sometimes been fearful of stepping on sidewalk cracks. (T)
- 20. I have wondered whether the spirits of the dead can influence the living. (T)
- 23. Things sometimes seem to be in different places when I get home, even though no one has been there. (T)
- 24. I think I could learn to read other's minds if I wanted to. (T)
- 28. At times I perform certain little rituals to ward off negative influences. (T)
- 29. I have felt that I might cause something to happen just by thinking too much about it. (T)
- 34. I have occasionally had the silly feeling that a TV or radio broadcaster knew I was listening to him. (T)
- 39. I have sometimes felt that strangers were reading my mind. (T)
- 49. Some people can make me aware of them just by thinking about me. (T)

50. I have worried that people on other planets may be influencing what happens on earth. (T)
53. The hand motions that strangers make seem to influence me at times. (T)
57. People often behave so strangely that one wonders if they are part of an experiment. (T)
66. I have had the momentary feeling that someones's place has been taken by a look-alike. (T)
69. I have felt that there were messages for me in the way things were arranged, like in a store window. (T)
77. I have noticed sounds on my records that are not there at other times. (T)
84. At times I have felt that a professor's lecture was meant especially for me. (T)
106. I have sometimes had the passing thought that strangers are in love with me. (T)
109. I have had the momentary feeling that I might not be human. (T)
113. Horoscopes are right too often to be a coincidence. (T)
138. I have sometimes sensed an evil presence around me, although I could not see it. (T)
139. If reincarnation were true, it would explain some unusual experiences I have had. (T)
141. The government refuses to tell us the truth about flying saucers. (T)
55. Numbers like 13 and 7 have no special powers. (F)
60. I have never had the feeling that certain thoughts of mine really belonged to someone else. (F)
65. It is not possible to harm others merely by thinking bad thoughts about them. (F)
71. When introduced to strangers, I rarely wonder whether I have know them before. (F)
91. I almost never dream about things before they happen. (F)
116. Good luck charms don't work. (F)
140. I have never doubted that my dreams are the products of my own mind. (F)

Physical Anhedonia Scale

6. There just are not many things that I have ever really enjoyed doing. (T)
8. The sound of rustling leaves has never much pleased me. (T)
10. I have always hated the feeling of exhaustion that comes from vigorous activity. (T)
12. I don't understand why people enjoy looking at the

- stars at night. (T)
16. The color that things are painted has seldom mattered to me. (T)
22. The sounds of a parade have never excited me. (T)
25. The beauty of sunsets is greatly overrated. (T)
41. Sex is okay, but not as much fun as most people claim it is. (T)
44. Flowers aren't as beautiful as many people claim. (T)
47. Poets always exaggerate the beauty and joys of nature. (T)
52. I usually finish my bath or shower as quickly as possible just to get it over with. (T)
56. I have seldom cared to sing in the shower. (T)
62. I've never cared much about the texture of food. (T)
72. I never wanted to go on any of the rides at an amusement park. (T)
75. I have never found a thunderstorm exhilarating. (T)
79. I never have the desire to take off my shoes and walk through a puddle barefoot. (T)
89. I think that flying a kite is really silly. (T)
93. I have had very little fun from physical activities like walking, swimming, or sports. (T)
101. One food tastes as good as another to me. (T)
103. I have had very little desire to try new kinds of foods. (T)
105. I have always found organ music dull and unexciting. (T)
108. I have seldom enjoyed any kind of sexual experience. (T)
112. I don't know why some people are so interested in music. (T)
115. I have usually found soft music boring rather than relaxing. (T)
119. The smell of dinner cooking has hardly ever aroused my appetite. (T)
121. I have often felt uncomfortable when my friends touch me. (T)
122. Dancing, or the idea of it, has always seemed dull to me. (T)
123. Sunbathing isn't really more fun than lying down indoors. (T)
134. The warmth of an open fireplace hasn't especially soothed and calmed me. (T)
135. On hearing a good song I have seldom wanted to sing along with it. (T)
142. I've never cared to sunbathe; it just makes me hot. (T)
2. I have sometimes enjoyed feeling the strength of my muscles. (F)

- On seeing a soft, thick, carpet, I have sometimes had the impulse to take off my shoes and walk barefoot on it. (F)
13. I have been fascinated with the dancing of flames in a fireplace. (F)
15. I have often enjoyed receiving a strong, warm handshake. (F)
18. The taste of food has always been important to me. (F)
19. I have always loved having my back massaged. (F)
21. The bright lights of the city are exciting to look at. (F)
27. When I have seen a statue I have had the urge to feel it. (F)
31. After a busy day, a slow walk has often felt relaxing. (F)
33. I have always had a number of favorite foods. (F)
37. It has always made me feel good when someone I care about reaches out to touch me. (F)
43. When I have walked by a bakery, the smell of fresh bread has often made me hungry. (F)
45. It has often felt good to massage my muscles when they are tired or sore. (F)
63. When I pass by flowers, I have often stopped to smell them. (F)
68. I like playing with and petting soft little kittens or puppies. (F)
70. Beautiful scenery has been a great delight to me. (F)
73. I have sometimes danced by myself just to feel my body move with the music. (F)
74. I have often found walks to be relaxing and enjoyable. (F)
82. When eating a favorite food, I have often tried to eat slowly to make it last longer. (F)
90. I have usually found lovemaking to be intensely pleasurable. (F)
94. A good soap lather when I'm bathing has sometimes soothed and refreshed me. (F)
98. The first winter snowfall has often looked pretty to me. (F)
100. When I'm feeling little sad, singing has often made me feel happier. (F)
110. Sex is the most intensely enjoyable thing in life. (F)
117. Standing on a high place and looking out over the view is very exciting. (F)
125. Trying new foods is something I have always enjoyed. (F)
127. The sound of organ music has often thrilled me. (F)
129. The sound of the rain falling on the roof has made me feel snug and secure. (F)

137. I have often enjoyed the feel of silk, velvet, or fur. (F)
143. A brisk walk has sometimes made me feel good all over. (F)

Impulsive Nonconformity Scale

78. When I start out in the evening, I seldom know what I'll end up doing. (T)
144. I often get so mad that I lose track of some of the things I say. (T)
146. Thinking things over too carefully can destroy half the fun of doing them. (T)
148. I usually quit before finishing one activity in order to start something else. (T)
149. As often as once a month I have become so angry that I have had to hit something or someone to relieve my anger. (T)
150. I frequently overeat and wonder why later. (T)
151. Most people say "please" and "thank you" more often than is necessary. (T)
153. When I want something, delays are unbearable. (T)
154. I don't have much sympathy for people whom I can push around and manipulate easily. (T)
155. Most of the mourners at funerals are just pretending to be sad. (T)
156. My way of doing things is apt to be misunderstood by others. (T)
157. Most people think of me as reckless. (T)
158. I always let people know how I feel about them, even if it hurts them a little. (T)
159. I almost always do what makes me happy now, even at the expense of some distant goal. (T)
160. I have had to invent some good excuses to get out of work or taking exams. (T)
161. I think people spend too much time safeguarding their future with savings and insurance. (T)
162. I break rules just for the hell of it. (T)
163. I usually find myself doing things "on impulse." (T)
164. I usually act first and ask questions later. (T)
166. I prefer being spontaneous rather than planning ahead. (T)
168. I sometimes do dangerous things just for the thrill of it. (T)
169. No one seems to understand me. (T)
170. I let go and yell a lot when I'm mad. (T)
171. I find it difficult to remain composed when I get into an argument. (T)

172. Long-term goals are not as important for me as living for today. (T)
173. During one period when I was a youngster I engaged in petty thievery. (T)
175. I often do unusual things just to be different from other people. (T)
178. In school I sometimes got in trouble for cutting up. (T)
180. I like to use obscene language to shock people. (T)
181. People who drive carefully annoy me. (T)
183. I liked to annoy my high school teachers. (T)
184. When I really want something, I don't care how much it costs. (T)
186. My parents often objected to the kind of people I went around with. (T)
187. I would probably purchase stolen merchandise if I knew it was safe. (T)
189. I do many things that seem strange to others but don't seem strange to me. (T)
190. I wouldn't worry if my bills were overdue. (T)
192. I usually laugh out loud at clumsy people. (T)
87. It worries me if I know there are mistakes in my work. (F)
145. I never get so angry that I can't speak coherently. (F)
147. It's important to save money. (F)
152. My friends consider me a cool and controlled person. (F)
165. I rarely act on impulse. (F)
167. I always stop at red lights. (F)
176. I usually consider different viewpoints before making a decision. (F)
179. Being in debt would worry me. (F)
182. If I burped loudly while having dinner at the house of someone I knew, I would be embarrassed. (F)
188. I have never been in trouble with the law. (F)
191. I try to remember to send people birthday cards. (F)
194. I avoid trouble whenever I can. (F)
195. It would embarrass me a lot to have to spend a night in jail. (F)
198. I usually control my feelings well. (F)

Appendix G

Correlations

Table 5 Extended.

Correlations between the Perceptual Abberation Scale (Per. Ab.), the Magical Ideation Scale (Mag. Id.), the Physical Anhedonia Scale (Phy. An.), and the Impulsive Nonconformity Scale (Imp. Non.).

	Per. Ab.	Mag. Id.	Phy. An.
Mag. Id.	.688		
Phy. An.	-.174	-.175	
Imp. Non.	.349	.348	.227

(N = 223, 5 observations deleted)*

	Per. Ab.	Mag. Id.	Phy. An.
Mag. Id.	.689		
Phy. An.	-.175	-.178	
Imp. Non.	.349	.352	.228

(N = 221, 7 observations deleted)*

	Per. Ab.	Mag. Id.	Phy. An.
Mag. Id.	.677		
Phy. An.	-.167	-.166	
Imp. Non.	.386	.373	.218

(N = 222, 6 observations deleted)*

	Per. Ab.	Mag. Id.	Phy. An.
Mag. Id.	.679		
Phy. An.	-.167	-.174	
Imp. Non.	.330	.339	.230

(N = 222, 6 observations deleted)*

* Deletion was on the basis of a standardized residual greater than 3.0, or on a single observation accounting for more than 2% of the variance in the dependent measure.