7-1964

Construction and Validation of Six Mental Health Sub-Scales by Age, Sex, and Educational Status

Thomas James Boynton

CONSTRUCTION AND VALIDATION OF SIX MENTAL HEALTH SUB-SCALES BY AGE, SEX, AND EDUCATIONAL STATUS

by

Thomas James Boynton

A thesis presented to the
Faculty of the School of Graduate
Studies in partial fulfillment
of the
Degree of Master of Arts

Western Michigan University
Kalamazoo, Michigan
July 1964

ACKNOWLEDGMENTS

The author wishes to express his sincere gratitude for the amount of time and effort given by Dr. Jerome G. Manis as the committee chairman of this thesis. His thorough knowledge of research methodology and social-psychological theory provided a basis by which the findings were able to take on a more meaningful character than they would have had otherwise. The author also wishes to thank Dr. Robert Jack Smith who served on the thesis committee for the many valuable suggestions regarding the nature of mental institutions as well as for a number of anthropological insights into the mental health problem. A good deal of the credit for insights gained by this author through the analysis should also go to Dr. David M. Lewis who provided a number of helpful suggestions pertaining to the theory of mental illness. The co-operation of Mr. Jack R. Meagher, Director of the Western Michigan University Computer Center, is also greatly appreciated. Finally, the author wishes to thank the staff at the Kalamazoo State Hospital for their helpfulness during the data gathering.

Without these contributions, this analysis would not have been possible.

Thomas James Boynton

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES

Chapter I

INTRODUCTION

Purpose of the Study

The topic considered in this thesis is the construction and validation of six mental health sub-scales which are based upon procedures controlling for age, sex, and educational status. Each item in a 22 item mental health scale is tested for validity by age, sex, and educational status. Six sub-scales, one for each dichotomy of the three social groups, are constructed from this procedure. The validity of these sub-scales is then tested via six different criteria of mental health. Finally, the validity of the six sub-scales, the original 22 item scale, and a related 9 item scale[1] are compared by examining their strengths of relationship with the six mental health criteria.

The 22 Item Mental Health Scale, which is the basic scale of the study, consists of 22 verbal statements of symptoms judged to be

---

[1] A 9 item mental health scale was constructed by Manis et al., from the 22 Item Mental Health Scale. Both scales are described more thoroughly in Chapter II of this thesis. See Jerome G. Manis, Milton J. Brawer, Chester L. Hunt, and Leonard C. Kercher, "Validating a Mental Health Scale," American Sociological Review, Vol. 28, No. 1, February, 1963, pp. 113-116.

indicative of mental illness.[2] The 22 items were combined to form a summated scale and this scale was validated via the "known groups" and "independent criteria" techniques. The 22 Item Mental Health Scale was judged to be a reasonably valid "measure of group mental health."[3]

One limitation appears to reduce the generalization potential of the prior validation study, however. The 22 item scale was validated for entire samples without considering the behavioral characteristics of social sub-groups making up the samples. It is quite possible that these behavioral variations, which may include attitudes and values as well as more overt behaviors, are contaminating the results

---

[2]The items in the 22 Item Mental Health Scale were part of the total protocol used in the New York Midtown study. See Leo Srole, Thomas S. Langer, Stanley T. Michael, Marvin K. Opler, and Thomas A. C. Rennie, Mental Health in the Metropolis: The Midtown Manhatten Study, Vol. 1, New York: McGraw-Hill Book Co., Inc., 1962, Chapter IV and Appendices E and F. Responses to the 22 items "had correlated .4 or higher with the mental health ratings given the respondents by the Midtown study psychiatrists." (Jerome G. Manis, et al., Ibid., p. 108).

[3]The 22 Item Mental Health Scale had been originally designed to measure the mental health of individuals. The results of the validation study by Manis, et al., were essentially negative for individual differences, however. "As a measure of group differences, the mental health scales (referring to both the 22 and 9 item scales) appear more valid. In nearly every comparison, the mean scores of groups were in the hypothesized directions." "Mean scores on the scales do provide a rough index of a social phenomenon - group mental health." (Manis, et al., Ibid., p. 116). The present analysis treats the various

of the validation study. In other words, each item in the 22 Item Mental Health Scale may be measuring a number of factors other than mental health. These "other factors" may be associated with specific roles and/or general characteristics of certain social sub-groups. Whether or not these characteristics are indicative of mental health levels is open to question. Several authors, for example, have criticized the various operational definitions of mental illness on the grounds that they are biased by middle class value judgments on the part of the investigators. It has been suggested by the critics that the current definitions of mental illness simply do not apply to lower class culture.[4] Apparently, many measures of mental illness are not equally valid for the social classes.

---

mental health scales in the same manner, i.e., as measures of group behavior. Mean scores and proportions of persons falling between certain ranges of scores are used as indications of group mental health in this analysis.

[4] Among the numerous criticisms of mental illness definitions are the following: Kingsley Davis, "Mental Hygiene and Class Structure," Psychiatry, Vol. 1, 1938, pp. 55-65; Orville Gursslin, Raymond Hunt and Jack Roach, "Social Class, Mental Hygiene and Psychiatric Practice," Social Service Review, Vol. 33, September, 1959, pp. 237-244; Marie Jahoda, "Toward a Social Psychology of Mental Health," in R. Kotinsky and H. Witmer, Community Programs for Mental Health, Cambridge, Massachusetts: Harvard University Press, 1955; Sol W. Ginsburg, "The Mental Health Movement and Its Assumptions," in Kotinsky and Witmer (Ibid.,); Joseph Eaton, "The Assessment of Mental Health," American Journal of Psychiatry, Vol. 108, pp. 81-90; Allister Miles MacMillan, "A Survey Technique for Estimating Prevalence of Psychoneurotic and Related Types of Disorders in Communities," and

The same sort of objection may be raised regarding the validity of such scales for the sexes and age groups. This question is particularly relevant for those mental health measures which include a large number of items pertaining to the physical condition of the respondent. Given the fact that the aged, female, and low social status segments of our society exhibit the highest general morbidity rates,[5] one would logically expect such persons to report more physical disturbances than the young, male, and high social status groups. The question arises, however, as to whether or not such items are equally indicative of mental illness for these different social groups.[6]

---

also the discussions of this article in Benjamin Pasamanick, Epidemiology of Mental Disorders, Washington, D. C., American Association for the Advancement of Science, 1959; John A. Clausen, "Social Science Research in the National Mental Health Program," American Sociological Review, Vol. 15, No. 3, June, 1950, pp. 402-409; Alexander H. Leighton, "The Stirling County Study: Some Notes on Concepts and Methods" in Paul H. Hoch and Joseph Zubin, Comparative Epidemiology of Mental Disorders, New York: Grune and Stratten, 1961; John A. Clausen, "The Sociology of Mental Illness," in Robert K. Merton, Leonard Broom and Leonard S. Cottrell, Jr., Sociology Today, New York: Basic Books, Inc., 1962, and John A. Clausen, Sociology and the Field of Mental Health, New York: Russel Sage Foundation, 1956, pp. 16-18.

[5] Commission on Chronic Illness, Chronic Illness in a Large City, Vol. IV, Cambridge, Massachusetts, Harvard University Press, 1957, p. 50.

[6] For a discussion on the question of the validity of psychosomatic items for various social aggregates see Louisa P. Howe, "Problems in the Evaluation of Mental Health Programs," in Kotinsky and Witmer, op. cit., p. 287.

Thus, although the previous validation study demonstrates that the 22 Item Mental Health Scale is a valid instrument for estimating the prevalence of mental illness within and between relatively heterogeneous samples, whether or not the scale and its individual items are equally valid for age, sex, and educational status subgroups within the samples is a question which has been left unanswered.

Previous Literature

Research in the epidemiology of mental illness has produced a tremendous body of literature over the past decades.[7] Most of the studies have relied upon medical records of patients as the basic ingredient in their definitions of mental illness, however. This approach has been criticized by several authors on the grounds that

---

[7] A rough indication of the amount of literature in the mental health area can be seen in the pamphlet Current Sociological Research, New York: American Sociological Association, February, 1963, pp. 46-48, in which some 74 projects dealing with mental health are currently being conducted. The Sociological Abstracts and the Psychological Abstracts provide an even more comprehensive listing of studies in this area. As noted in the validation article by Manis, et al., op. cit., p. 108, there are very few validation studies in sociology. This topic is often given only brief notice in sociology textbooks and is generally pursued more systematically by psychologists concerned with individual test behavior.

it is limited by diagnostic inconsistencies and the ommission of untreated cases.[8]

An alternative technique for the epidemiological study of mental illness is the community survey in which mental health scales, such as the one under consideration in this study, are used to estimate the incidence and/or prevalence of mental illness within and/or between communities. Although this method is generally thought to have limited diagnostic capabilities, it does have the advantage of obtaining a more accurate approximation of the total count of disturbed persons within a given segment of the population.[9] The community survey is thus becoming an increasingly popular method for the sociological study of mental illness.[10]

Nearly all of the sociological studies dealing with mental illness have investigated the relationships between it and other social characteristics of their samples. A large number have dealt with mental illness variations between age groups, the sexes, and

---

[8] Jerome G. Manis, Milton J. Brawer, Chester L. Hunt, and Leonard C. Kercher, "Estimating the Prevalence of Mental Illness," American Sociological Review, Vol. 29, No. 1, February 1964, pp. 84-85.

[9] John A. Clausen,"The Sociology of Mental Illness," op. cit., p. 494.

[10] The increasing popularity of survey techniques for the study of mental illness in communities probably stems from the recognized limitations of using only treated cases, combined with the relatively expensive use of psychiatrists for examining and/or diagnosing entire samples. See John A. Clausen, Ibid., pp. 492-493.

and social classes. Although the findings are not always in agreement,[11] the following generalizations appear to be fairly consistent:

1. Mental illness tends to increase with age.[12]

2. Mental illness is more prevalent among females than males.[13]

---

[11] Manis, et al., "Prevalence Study," op. cit., p. 88, Footnote 25, note that "the relation between mental illness and age, sex and status is still controversial." They go on to point out that theoretical and methodological differences in the case definitions and locating techniques probably account for most of the differences in reported rates. This author feels that this situation may also contribute to differences between the studies findings regarding mental illness and age, sex, and status. It will be remembered that this is the very problem which the thesis seeks to solve.

[12] Cottrell has noted that the degree of adjustment to roles which our society assigns to its age-sex categories varies directly with the clarity with which such roles are defined. (See Leonard S. Cottrell, Jr.), "The Adjustment of the Individual to His Age and Sex Roles," in Theodore Newcomb, Eugene Hartley, et al., Readings in Social Psychology, New York: Holt, 1947, pp. 370-373. This generalization would seem to be particularly relevant in the study of mental health of these aggregates. There is ample evidence in the literature to show that the aged are more apt to be mentally ill. See: Leo Srole, et al., Midtown Study, op. cit., pp. 186 and 250; H. Warren Dunham, Sociological Theory and Mental Disorder, Detroit: Wayne State University Press, 1959, p. 241; Ernest M. Gruenberg, "A Mental Health Survey of Older Persons," in Paul H. Hoch and Joseph Zubin, op. cit., p. 21, Figure 4; E. Gartly Jaco, The Social Epidemiology of Mental Disorders, New York: Russel Sage Foundation, 1960, p. 32; Gerald Guring, Joseph Veroff and Shelia Feld, Americans View Their Mental Health, New York: Basic Books, 1960, pp. 189-170, and Commission of Chronic Illness, op. cit., p. 97.

[13] The statement regarding the clarity of role definitions and adjustment, (See Footnote 11 above), is also readily applicable to females. The following references support the contention that women tend to be more mentally ill than men: E. Gartly Jaco, Ibid., p. 96; Gerald Guring, et al., Ibid., p. 208-210 and Commission on Chronic Illness, Ibid., p. 97.

3. Mental illness is more prevalent in the lower social status groups.[14]

Much of the literature also reports variations in symptoms between these social groups. These variations depend, of course, upon the instrument used to measure the symptoms. In most cases, psychiatric diagnostic nomenclature is used to describe the symptoms.[15] Other criteria which are assumed to be indicative of abnormal symptoms are based upon a wide variety of theories dealing with personality.[16]

The findings pertaining to variations between various social groups in the prevalence of physical and non-physical symptoms are

---

[14] See for example: E. Gartly Jaco, Ibid., pp. 128-131 and 152-154; Gerald Guring, Ibid., pp. 189-197; August B. Hollingshead and Frederick C. Redlich, "Social Stratification and Psychiatric Disorders," American Sociological Review, Vol. 18, No. 2, April, 1953, pp. 163-169 and Bert Kaplan, Robert B. Reed, and Wyman Richardson, "A Comparison of the Incidence of Hospitalized and Non-Hospitalized Cases of Psychosis in Two Communities," American Sociological Review, Vol. 21, No. 4, August, 1956, pp. 472-479.

[15] In spite of the unreliability of application of the standard diagnostic nomenclature, there are those who feel that these categories provide the results of a great deal of clinical experience. (See John A. Clausen, "The Sociology of Mental Illness," op. cit., p. 493.) It is probably for this reason that these diagnosis categories continue to be used in studies of mental illness to such a degree.

[16] For an excellent and up-to-date discussion of theories on personality see Joseph M. Wepman and Ralph W. Heine, eds., Concepts of Personality, Chicago: Aldine Publishing Company, 1963.

especially pertinent to the present study.[17] Several studies have found that the aged, female, and lowest social status groups tend to have the highest prevalence of physical or psycho-somatic disturbances.[18]

As indicated earlier, however, the above relationships are subject to some question due to the apparent lack of reliable and valid measures of mental health.[19] Yet, these criticisms of the various measures of mental health have been somewhat impressionistic, presenting little or no empirical support for their arguments.

Most of the literature is thus marginal to this study, while that which appears to be directly pertinent offers little in the way of reporting systematic investigation of the problem posed in this thesis.

---

[17] The 22 Item Mental Health Scale contains 13 items which are considered to fall into the physical disturbance classification. These items are more thoroughly discussed in Chapter II.

[18] Perhaps the most pertinent example of this finding is presented by Gerald Guring, et al., op. cit., pp. 189-197. This study used a twenty item scale which is very similar to the 22 Item Mental Health Scale. See Allister Mile MacMillan, "The Health Opinion Survey: Technique for Estimating Prevalence of Psychoneurotic and Related Types of Disorders in Communities," Psychological Reports, Vol. 3, p. 325.

[19] See Footnote 4, above.

Hypotheses

Although the present analysis is somewhat exploratory in nature, two general hypotheses can be drawn up in accordance with the purpose of the study as well as with insights gained from the previous literature.

Hypothesis I: It is predicted that the prevalence of symptoms, as measured by the items in the 22 Item Mental Health Scale, will vary by age, sex, and educational status, the items pertaining to physical disturbances proving least valid for high age, female, and low educational status groups.

Hypothesis II: The six mental health sub-scales, in which variations in validity by age, sex, and educational status have been controlled, will prove to be more valid than the 22 and 9 item scales.

Chapter II

SCOPE AND METHODS

A number of concepts and procedures have been mentioned or implied in the previous chapter which have yet to be specifically defined. This is the task of the present chapter. The nature of the samples used in the study are presented, followed by descriptions of the data pertaining to age, sex, and educational status. The 22 and 9 item scales are presented and the physical and non-physical items of the scales are specified. Finally, the known groups item analysis, sub-scale construction, sub-scale validation, and validity comparisons procedures are briefly discussed.

Attention is focused first upon the three samples.

The Samples

Hospital receiving ward sample

The receiving ward sample was selected from a large State hospital for the mentally ill, located in Kalamazoo City, Michigan. The patients in the receiving ward were reported as newly admitted, being hospitalized for mental illness presumably severe enough to

"incapacitate them from functioning in their normal lives."[1] The patients were selected by a systematic sampling from complete listings of the patient population. Between 1/3 and 2/3 of the patients in this ward were privately interviewed by two researchers from a schedule which includes the 22 Item Mental Health Scale. Information pertaining to the age, sex, and independent estimates of mental health[2] were also obtained for each patient at this time. The educational data for the patients was collected by this author, after the original schedule had been administered.[3] Nine of the 110 completed schedules were judged to be invalid[4] leaving a total of 101 usable schedules for this sample.

Plainwell sample

Plainwell is the name of a relatively small community located

---

[1] Jerome G. Manis, Milton J. Brawer, Chester L. Hunt, and Leonard C. Kercher, "Validating A Mental Health Scale," American Sociological Review, Vol. 28, No. 1, February, 1963, p. 110.

[2] These data are described later in the chapter and also in Appendix A, sections I, II and III.

[3] The educational data for the hospital receiving ward patients was collected by this author in October, 1963, whereas the original data for the larger mental health study was collected in 1959.

[4] The judgments as to the validity of these schedules were based upon the interviewers' estimates of the patients' ability to comprehend and/or answer the questions in a coherent manner.

approximately 10 miles north-west of Kalamazoo, Michigan. The Plainwell sample was drawn by a systematic sampling of the city directory. Every seventh dwelling unit was contacted at least twice, and of the 130 contacted units, 8 were refusals, 7 were incorrect listings and 31 were not at home.[5] The remaining 84 respondents were interviewed in their homes from schedules containing the 22 Item Mental Health Scale, as well as items referring to age, sex, and educational status.

Kalamazoo sample

Kalamazoo County lies approximately half way between Detroit and Chicago, a number of major highways and railroads intersecting its borders. The city of Kalamazoo, which contributes heavily to the Kalamazoo sample, has a number of diversified industries, small businesses and educational and research institutions.[6]

The sample was drawn from Kalamazoo county via a two-stage design by Leslie Kish and Bernard Lazerwitz of the University of Michigan Survey Research Center. "From a three strata 8% master sample of

---

[5] Manis, et al., "Validation Study," op. cit., p. 109, Footnote 7.

[6] Jerome G. Manis, Milton J. Brawer, Chester L. Hunt, and Leonard C. Kercher, "Estimating the Prevalence of Mental Illness," American Sociological Review, Vol. 29, No. 1, February 1, 1964, p. 85.

dwelling units, a one-third systematic sub-sample was drawn from each stratum. Of the original 1,361 addresses, 42 were unoccupied, 18 were not residential and 10 could not be located. 1,293 house households were actually contacted but 48 were not at home (after 4 calls), 53 were refusals and 9 schedules were incomplete."[7] The 1,183 usable schedules from this sample are almost identical to those used for the Plainwell sample.[8] The Kalamazoo respondents were interviewed in their homes and data pertaining to their age, sex, and educational status, as well as their responses to the 22 Item Mental Health Scale were gathered at this time.

Age, Sex, and Educational Status

The operational specifications for age, sex, and educational status, although somewhat obvious, are presented here so as to make their meanings completely clear.

The age of each respondent was obtained at the time of the interview for all three samples. The respondent's age at his last birthday is the measure of this variable.

The sex of each respondent was also recorded by the interviewer

---

[7] Manis, et al., "Validation Study," op. cit., p. 109 Footnote 7.

[8] The Plainwell sample was used by Manis, et al., in the pre-test of the schedule used for the Kalamazoo data.

at the time of the interview.

The educational status of each respondent was determined by his last grade completed in school. Educational attainment has been used by a number of authors as an index of social stratification. Since many students of social stratification feel that social class or status, has many dimensions, however, the term "educational status" will be used throughout the thesis.[9]

Although a number of dimensions, or indicators of social class could have been used in the study, educational status was chosen because it was, first, the most readily obtainable measure, and secondly, because it has certain advantages over other indicators of social class. Occupational status, for example, has been cited as the best single indicator of social status,[10] yet, there are several

---

[9]The term "educational status" is used in this analysis in order to clearly differentiate it from the other dimensions of social class. For descriptions and discussions of the many dimensions of social class see: Max Weber, "Class, Status, Party," Essays in Sociology, translated by H. H. Gerth and C. Wright Mills, New York: Oxford University Press, Inc., 1946, pp. 180-195; Bernard Barber, Social Stratification, Harcourt, Brace and Company, New York: 1957, chpts. 1-8; Joseph A. Kahl, The American Class Structure, New York: Holt, Rinehart and Winston, 1961; and Reinhard Bendix and Symore M. Lipset, Class, Status and Power, Free Press of Glencoe, 1963, chpt. I.

[10]Joseph A. Kahl and James A. Davis, "A Comparison of Indexes of Socio-Economic Status," American Sociological Review, Vol. 20, pp. 317-325; Paul K. Hatt, "Occupational and Social Stratification," American Journal of Sociology, Vol. 55, pp. 533-543; and Bernard Barber, op. cit., pp. 184-186.

objections to the use of this measure. Perhaps the most outstanding limitation of occupational status categories is that they do not, in and of themselves, form an ordinal level of measurement. Researchers have noted that occupational status categories, the status of which must be empirically defined, shift their positions within the social status hierarchy from time to time and from place to place.[11] This shift in status need not occur all along the continuum, but may occur between only a few categories, reversing their positions in such a manner as to destroy the ordinal effect of the measure.

Educational status categories are not likely to manifest the same characteristics. Although the prestige of a given educational category may change from time to time or place to place, the shift tends to occur all along the continuum. Educational status categories, by their very nature, form an ordinal scale.[12] In other words, it is unlikely that a 12th grade education will be judged to be inferior to a 10th grade education, a 9th grade education superior to a 10th

---

[11] See eg., National Opinion Research Center, "Jobs and Occupations: A Popular Evaluation," Opinion News, Vol. 9, pp. 3-13. In this study, variations in job evaluation by section of the country are discussed. Manis discusses the possibility of the status of different occupational categories shifting from time to time in a preliminary unpublished report dealing with the relationship between mental health and social status. Jerome G. Manis, Center for Sociological Research, Western Michigan University, Kalamazoo, Michigan.

[12] For a description of different levels of measurement see,

grade education, etc. These categories are more apt to maintain their ordinal characteristic than those pertaining to occupational status.

The age, sex, and educational status variables are dichotomized in the present study in order to facilitate the analysis.[13] The ages of the respondents are divided into 44 years or less, and 45 years and older. Sex is, of course, divided into males and females. Persons having completed 11 years of schooling, or less constitute the low educational status category, while those having a 12th grade education or more make up the high educational status group.

Mental Health Scales

As noted in Chapter I, the study has a number of measures designed to indicate the relative levels of mental illness among various social aggregates. Descriptions of these scales are given below.

The 22 Item Mental Health Scale

The basic mental health measure of the study is the 22 Item Mental Health Scale. The items making up this scale were a part of

---

Hubert M. Blalock, Jr., Social Statistics, New York: McGraw-Hill Book Company, Inc., 1960, pp. 11-16.

[13] These data were dichotomized because of the relatively small size of the hospital and Plainwell samples. This procedure also facilitates the use of the two sample test of proportions. See footnote 25.

the total protocol used in the Midtown Manhatten study.[14] The items in the Midtown schedule were constructed by the study psychiatrists on the basis of their clinical experience. The 22 items making up the scale under consideration here were chosen on the basis of their agreement with ratings given the Midtown respondents by the psychiatrists; only those items which correlate .4 or higher with the independent ratings of the psychiatrists were included. The 22 items were formed into a simple additive scale in which a score of "0" indicates relatively good mental health, while higher scores denote progressively poorer mental health. The 22 Item Mental Health Scale was later validated and found to be a reasonably accurate measure of group mental health.[15]

The 22 Item Mental Health Scale has a large number of items indicating physical disturbances. Since variations by physical and non-physical symptoms are examined in the study, a classification of the 22 items, according to these categories, is presented in Figure I.

---

[14] Leo Srole, Thomas S. Langer, Stanley T. Michael, Marvin K. Opler, and Thomas A. C. Rennie, Mental Health in the Metropolis: The Midtown Manhatten Study, Vol. 1, New York: McGraw-Hill Book Company, Inc., 1962, Chapter IV and Appendices E and F.

[15] Manis, et al., "Validation Study," op.cit., p. 116.

Figure 1

Individual Items of the 22 Item Mental Health Scale Classified by Physical and Non-physical Nature of the Disturbance[16]

Items Indicating Physical Disturbances

1. Are you ever troubled with headaches?
2. Do you ever have any trouble getting to sleep or staying asleep?
3. Do your hands ever tremble enough to bother you?
4. Have you ever been bothered by shortness of breath when you were not exercising or working hard?
5. Have you ever been bothered by "cold sweats?"
6. Have you ever been bothered by your heart beating hard?
7. Have you ever had fainting spells?
8. How would you describe your appetite?
9. I feel weak all over much of the time.
10. I am bothered by acid (sour) stomach several times a week.
11. Every so often I suddenly feel hot all over.
12. There seems to be a fullness (clogging) in my head or nose much of the time.
13. I have personal worries that get me down physically.

Items Indicating Non-physical Disturbances

1. Are you ever bothered by nervousness (irritable, fidgety, tense)?
2. In general, would you say that you are in very low spirits most of the time?
3. I am the worrying type.
4. I feel somewhat apart even among friends.
5. I have periods of such restlessness that I cannot sit long in a chair.

---

[16] Items 2 and 6 in the non-physical disturbance category are reworded for purposes of clarity. See Appendix A, section IV.

6. My memory does not seem to be all right.
7. I have had periods of days, weeks, or months when I couldn't get going.
8. Nothing ever turns out for me the way I want it to.
9. I sometimes can't help wondering whether anything is worthwhile anymore.

The items above were categorized by this author on the basis of their apparent face validity. Most of the items appear to be clearly tapping either physical or non-physical conditions of the respondents. Not all the items are equally obvious in this respect, however. The item referring to worries getting the respondent down physically appears to be measuring a condition in which a psychological condition results in a physical malady. Whether or not this item should be classified as a strictly physical symptom is thus open to question, although to categorize the item as a non-physical condition would be equally questionable. The item referring to the respondent's inability to sit in a chair due to restlessness might be considered a physical symptom, if one views this condition strictly in terms of the neuro-muscular system. The classification is thus somewhat arbitrary in a few cases. It is felt that this situation introduces no serious limitation to the study however, since the items are analyzed separately in the sub-scale construction procedures.[17]

---

[17] These procedures are described in the next section and also in Chapter III.

The 9 Item Mental Health Scale

A 9 item mental health scale was constructed from the items in the 22 item scale by Manis, et al.[18] The items for the 9 item scale were selected on the basis of their ability to discriminate between the hospital receiving ward patients and three other samples - hospital predischarge ward patients, Plainwell respondents, and Kalamazoo respondents.[19] The 9 item scale was validated in the same manner as was the 22 item scale. The items included in the 9 Item Mental Health Scale are presented in Figure 2.

Figure 2

Individual Items of the 9 Item Mental Health Scale[20]

1. Are you ever bothered by nervousness (irritable, fidgety, tense)?
2. In general, would you say that you are in very low spirits most of the time?
3. I am the worrying type.
4. I feel weak all over much of the time.
5. I have periods of such restlessness that I cannot sit long in a chair.

---

[18] Manis, et al.,"Validation Study," op. cit., pp. 114-116.

[19] A college student sample was excluded from the construction procedures of the 9 Item Mental Health Scale. (Ibid.).

[20] Item 2 is re-worded for purposes of clarity. See Appendix A, section IV.

6. I am bothered by acid (sour) stomach several times a week.
7. Nothing ever turns out for me the way I want it to.
8. I have personal worries that get me down physically.
9. I sometimes can't help wondering whether anything is worthwhile anymore.

The present study also has 6 mental health sub-scales which are formed from the items in the 22 Item **Mental Health** Scale. The construction of the sub-scales involves the known groups item analysis, so their nature is discussed in the next section which deals with this procedure.

Known Groups Item Analysis and Sub-Scale Construction

The known groups item analysis involves two samples, the degree of mental health of which is either known, or can be estimated. The receiving ward patient sample is classified as the "mentally ill" group, while the Plainwell sample is considered to be "mentally well."

The mental health of the receiving ward patients is perhaps the most accurately estimated. These persons had been hospitalized due to mental illness which had inhibited them from coping successfully with the frustrations of their environments. Moreover, they had not sufficiently recovered from their illness to be admitted to other

wards as chronically ill, or for predischarge therapy.[21]

By contrast, only 1.2% of the Plainwell respondents had ever been in a mental hospital for treatment, and only 3.6% had ever seen a psychologist, or a medical doctor, because of nervousness or mental problems.[22] Obviously, none of the Plainwell respondents were hospitalized at the time of the interview, and although a few may have been as sick, or sicker than persons in the receiving wards of the hospital, as a group, their mental health is probably better.[23]

Validity of the items in the 22 Item Mental Health Scale

The ability of each item in the 22 Item Mental Health Scale to discriminate between the "mentally ill" and the "mentally well" samples, described above, is determined in this thesis. In this procedure, the proportion of "positive" or "sick" responses to an item[24] by one sub-group in the receiving wards of the hospital is

---

[21] Manis, et al, "Validation Study," op. cit., p. 110.

[22] These figures were computed by this author from the Plainwell data.

[23] This is merely an assumption, however, and the operational specification of mental illness is ultimately based upon whether or not the persons making up the samples were hospitalized for mental illness at the time of the interviews.

[24] The sick, or positive responses to the 22 items are indicated by an asterisk in Appendix A, section IV.

compared to the proportion of sick responses to this item given by the comparable sub-group in the Plainwell sample. For example, the proportion of males in the receiving wards responding positively to an item is compared to the proportion of males in the Plainwell sample who give positive responses to the same item. This procedure is followed for the dichotomies of age, sex, and educational status for all 22 items. Those items which discriminate between the two samples for a given sub-group at the .01 level of significance are considered to be valid indicators of mental illness for that sub-group.[25]

The nature of variations in response to these valid items by age, sex and educational status sub-groups are examined and discussed.

Construction of the mental health sub-scales

The construction of the 6 mental health sub-scales is based upon the known group analysis. The selection of an item going into a sub-scale is determined by its ability to discriminate between the hospital receiving ward and the Plainwell samples by age, sex, and educational status, at the .01 level.

---

[25]The two sample significance test between proportions is used in the known group analysis.

The social characteristics - age, sex, and educational status have two sub-scales each. Items which discriminate for both dichotomies of a characteristic are included in both sub-scales.[26] In addition to these common items, each sub-scale includes items which discriminate for one dichotomy of the characteristic, but not for the other. For example, the sub-scale for males includes items which discriminate between the mentally ill and well samples for, both, males and females, and also items which discriminate for males only. The sub-scale for females, on the other hand, contains items which discriminate for both sexes, as well as items which discriminate for females only.

The six mental health sub-scales for the age, sex, and educational sub-groups are labelled as follows:

1. Low Age Mental Health Scale
2. High Age Mental Health Scale
3. Male Mental Health Scale
4. Female Mental Health Scale
5. Low Educational Status Mental Health Scale
6. High Educational Status Mental Health Scale

---

[26] Items discriminating for two or more social sub-groups could have been formed into a sub-scale, but because they are so few in number, this procedure is not likely to add much to the analysis. See Tables I, II and III of Chapter III.

Validation of the Mental Health Sub-Scales

The validity of the six sub-scales is tested by using six different criteria of mental health. These criteria are described below.

Sub-scale validation by known groups analysis

A known group analysis, very similar to the one used to construct the sub-scales, is used to test the validity of these scales. The difference between the two procedures lies in the samples used and the nature of the measures tested.

In the known groups item analysis, Plainwell is considered to be the mentally well group, while in the sub-scale known groups analysis, the Kalamazoo sample is defined as the mentally well group. The rationale for defining the Kalamazoo sample as mentally well is much the same as that used for the Plainwell sample. Although some of the Kalamazoo respondents may be as mentally ill, or more ill than some or all of the hospital patients, as a group, their mental health is probably better.[27] It should be noted that both procedures use the

---

[27] Approximately .5% of the Kalamazoo respondents reported being in a mental hospital for treatment, 1.9% reported seeing a psychologist about mental problems and 17.0% said that they had seen a medical doctor about nervousness or mental problems. As in the case with the Plainwell sample, the ultimate basis for defining the Kalamazoo sample as mentally well is that they are not hospitalized, however.

hospital sample as the mentally ill group.[28]

The second major difference between the procedures is that in the sub-scale construction the items are the basic unit of analysis, whereas in the sub-scale validation, the discriminatory ability of the sub-scales, as wholes, is the focus of analysis.

Sub-scale validation by independent criteria of mental health within the hospital sample[29]

Three independent criteria of the hospital patients' mental health levels are used to further test the validity of the six sub-scales.

The first criterion of mental health refers to the floor assignment of the patients at the hospital. It is the policy of the staff at the Kalamazoo State Hospital to assign patients to different floors according to the severity of their mental illness. The third floor receives the most severe mental cases, the second floor the next most severe, while the least disturbed patients are assigned to the first floor.

---

[28]The possible ramifications of this situation are discussed in Chapter V of this thesis.

[29]Data for the floor assignment and previous admissions criteria were taken directly from hospital files. The attendants' ratings were obtained by interviewers directly from the attendants.

The second criterion of mental health for the mental patients is the attendant's mental health ratings of each patient under his supervision along a five point scale ranging from "very poor" to "very good" mental health. For the purposes of this analysis, these ratings are grouped into: "good to very good," "fair," and "poor to very poor" categories.[30]

The third independent criterion of the patients' mental health levels is based upon the number of previous admissions of each patient to the hospital.

Sub-scale validation by independent criteria of mental health within the Kalamazoo sample

Two criteria of the mental health levels of the Kalamazoo respondents make up the final tests of validity for the sub-scales.

The first estimate of the mental health of the Kalamazoo sample is measured by an item inquiring as to whether or not the respondent "has been to see a doctor about mental problems or nervousness."

The second criterion refers to the respondent's score on a 45 item mental health scale constructed by Manis, et al., in conjunction

---

[30] These categories were trichotomized due to the small numbers involved in this sample.

with the consulting psychiatrists to the project.[31] The scores on the 45 item scale are dichotomized into the groups: scores zero to one; scores of two or higher. The higher scores denote progressively poorer mental health.

The sub-scales are treated as simple additive measures on which the respondents receive total scores. The respondents' scores on the appropriate sub-scales, i.e., according to their age, sex, and educational status, are compared with their mental health ratings as measured by the six criteria, described above. Statistical tests[32] are run in order to determine whether or not the scales are significantly related to the criteria in a positive manner.[33]

Comparisons of the Validity of the Sub-Scales, the 22 Item, and 9 Item Mental Health Scales

Comparisons between the validity of the six sub-scales, the 22

---

[31] Manis et al., "Validation Study," op. cit., pp. 113-116. The items of the 45 Item Mental Health Scale are presented in Appendix A, section VIII.

[32] Chi-square tests are used to determine the significance of the relationships. All tests are based upon 2 x 2 contingency data. See Appendix C.

[33] It is reasoned that if the sub-scales are valid measures of mental illness, then the respondents' scores on the scales should increase as their estimated mental health becomes worse. A positive or direct relationship between the scales and the estimates of mental health, as measured by the six criteria, is thus expected.

Item Mental Health Scale, and the 9 Item Mental Health Scale are made in the study by examining the strengths of relationship between the scales and the six mental health criteria described above. These comparisons are accomplished via the Pearsonian contingency coefficient.[34]

The first step in the series of operations, described above, involves the analysis of the items in the 22 Item Mental Health Scale and the subsequent sub-scale construction. A more thorough presentation of the rationale for these procedures as well as their results are discussed in the next chapter.

---

[34] For a description of this statistic see Hubert M. Blalock, Social Statistics, New York: McGraw-Hill Book Company, Inc., 1960, pp. 230-231.

Chapter III

KNOWN GROUPS ITEM ANALYSIS AND SUB-SCALE CONSTRUCTION

In this chapter, the validity of each item in the 22 Item Mental Health Scale is ascertained for age, sex, and educational status sub-groups. Six mental health sub-scales, one for each social sub-group, are constructed from these procedures. Before describing the results, a brief discussion of the concept of validity is presented in order to clarify its meaning, particularly in reference to the present analysis.

The Concept of Validity

There are a number of definitions of validity to be found in the various articles and books dealing with the subject. Most definitions are in agreement, usually differing only in the terms used to describe identical, or nearly identical characteristics of the concept.

A measuring instrument has validity when it measures that which it purports to measure. Stated another way, the "validity of a measuring instrument may be defined as the extent to which differences in scores on it reflect true differences among individuals, groups, or situations in the characteristic which it seeks to measure...rather

than constant or random errors."[1] The "pragmatic" approach to validity is concerned with the usefulness of the instrument as an indicator, or predictor of some behavior or characteristic of an individual or group. "What is essential in this approach to validation is that there be a reasonably valid and reliable criterion with which the scores on the measuring instrument can be compared."[2] The pragmatic approach to validity is used in the present analysis, the usefulness of the 22 items as indicators of mental illness for certain subgroups being the basic concern.[3]

There are several methods which can be used to test the pragmatic validity of items and/or scales.[4] In this study, the "known group" procedure is used to validate the items in the 22 Item Mental Health

---

[1] Claire Selltiz, Marie Jahoda, Morton Deutsch and Stuart W. Cook, Research Methods in Social Relations, New York: Holt, Rinehart and Winston, 1962, p. 155.

[2] Ibid., p. 157.

[3] The validity of the six sub-scales, as wholes, is tested in Chapter IV.

[4] See eg., Claire Selltiz, et al., op. cit., pp. 154-166; William J. Goode and Paul K. Hatt, Methods in Social Research, New York: McGraw-Hill Book Company, Inc., 1952, pp. 237-239, and Frank S. Freeman, Theory and Practice of Psychological Testing, New York:Henry Holt and Company, 1957, pp. 26-41.

Scale. The receiving ward patients are defined as the "mentally ill" group, while the Plainwell respondents are considered to be "mentally well." Although this classification is ultimately based upon the hospitalization status of the samples, some evidence was presented in Chapter II supporting the contention that the receiving ward patients are, in fact, more mentally ill than the Plainwell respondents.[5] These samples are thus judged to be "known groups," in terms of their mental illness. The validity of the 22 items is ascertained by testing each item's ability to discriminate between these two samples.

There is still one requirement of the definition of validity which has hereto been ignored. In the definition presented above, it is specified that the measuring instrument should not be tapping factors which do not measure what the instrument has been designed to measure. An item in the 22 Item Mental Health Scale may be measuring a number of factors other than mental illness. For example, the item inquiring as to whether or not the respondent has been bothered by his heart beating hard may be tapping a physical condition which has little or nothing to do with mental health. The prevalence of heart

---

[5] See page 23 of the thesis.

trouble is much greater for older persons.[6] They would thus be more apt to answer this item in the affirmative, even though they may have as good, or better mental health than younger persons. The item referring to the patient's inability to "get going" is another example of a possible source of error if used without taking into consideration the nature of the respondent's attitudes. Several authors have noted that lower class people are less inclined to be punctual and often exhibit a considerable amount of procrastination in their duties.[7] A lower class person, or one with low educational status, might thus be expected to answer this question positively, although he may have "good mental health."

Obviously, the above comments do not begin to exhaust the possible sources of constant errors within the 22 Item Mental Health Scale, but merely serve as an indication of the nature of the problem. Any number of extraneous factors, such as race, religion and ethnic

---

[6] Commission on Chronic Illness, Chronic Illness in a Large City, Vol. IV, Cambridge, Massachusetts, Harvard University Press, 1957, p. 50.

[7] See eg., Genevieve Knupfer, "Portrait of an Underdog," Public Opinion Quarterly, 11, 1947 (Spring), pp. 103 and 112-113; Orville Gursslin, Raymond Hunt and Jack Roach, "Social Class, Mental Hygiene and Psychiatric Practice," Social Service Review, 33, 1959 (September), p. 240, and Kingsley David, "Mental Hygiene and Class Structure," Psychiatry, 1, 1938, pp. 56-57.

background can introduce biases into the scale. An examination of all possible sources of error is, of course, far beyond the scope of this or any other study. The present known groups analysis thus focuses upon the validity of the individual items in the 22 Item Mental Health Scale by age, sex, and educational status. The results of this analysis are presented and discussed below.

Validity of the 22 Items by Age

The proportions of positive responses to the 22 Items were computed by age for the hospital receiving ward and Plainwell samples.[8] These proportions were tested for significance of differences. The resulting Z scores are presented in Table I.

Table I

Z Scores Describing Differences in Proportions of Sick Responses to the 22 Items Between Receiving Ward and Plainwell Samples by Age

Physical Disturbances

	Age	
	44 years or less	45 years or more
1. Headaches, often	.194	-.281

[8] The proportions corresponding to the Z scores in Tables I, II and III are presented in Appendix B, section IV.

Table I continued

	Age	
	44 years or less	45 years or more
2. Trouble sleeping, often	2.494 *	.322
3. Hands tremble, often	1.277	2.001
4. Shortness of Breath, often	2.098	.532
5. Cold sweats, often	2.041	.283
6. Heart beating hard, often	2.041	.053
7. Fainting spells, more than a few times	-.168	1.324
8. Appetite, poor	.898	1.228
9. Weak all over, yes	3.349	1.467
10. Acid Stomach, yes	1.151	2.526 *
11. Hot all over, yes	.929	.924
12. Fullness in head, yes	.850	.284
13. Worries get you down physically, yes	4.549 **	3.473 **
Non-Physical Disturbances		
1. Nervousness, often	2.011	2.005
2. Spirits, very low	1.572	2.570 *
3. Worrying type, yes	.971	2.864 *
4. Feel apart, yes	2.095	2.589 *
5. Restlessness, yes	3.373 **	2.194

Table I continued

	Age	
	44 years or less	45 years or more
6. Memory all right, no	1.649	1.914
7. Couldn't get going, yes	1.003	2.123
8. Nothing turns out, yes	2.636 *	3.650 **
9. Nothing worthwhile, yes	3.177 **	2.604 *

The Z scores were tested using one tail because the items were assumed to be tapping behavior most indicative of hospital patients.

* These Z scores are significant at the .01 level.
** These Z scores are significant at the .001 level.

The over-all trend described by the Z scores in Table I shows that the differences between the proportions of responses of the samples vary by age. Nine of the 13 Z scores for the physical items are largest for the younger group. The reverse situation is true for the non-physical items. In this case, six of the nine Z scores are larger for the older group. Although most of these differences are not significant at the .01 level, the trend remains essentially the same when only the significant Z scores are examined. Three physical items discriminate between the receiving ward and Plainwell samples for the younger group, while only two of these items discriminate for the older respondents. Five of the non-physical items prove

to be valid for the older group, but only three of the items in this category discriminate for the younger group.

In spite of the fact that the older respondents report physical disturbances more often than the young,[9] these items are less able to differentiate between the mentally ill and well samples for this age group. The aged, it will be recalled, tend to have higher morbidity rates and would thus be expected to report more physical disturbances. The above results clearly demonstrate, however, that these characteristics need not be indicative of mental illness. In other words, the results suggest that mental health scales, made up of a large number of physical items, may make the aged merely appear to be more mentally ill.

A total of six items discriminates between the known groups for the low age group, while seven items prove to be valid for the older group. The Low Age Mental Health Scale and the High Age Mental Health Scale are made up from these items, respectively.[10]

---

[9] See Appendix B, section IV.

[10] The proportions of sick responses to the items making up all six sub-scales are shown in Appendix B, section IV. The sub-scale mean scores are presented in Appendix B, section V.

Validity of the 22 Items by Sex

The Z scores describing the differences in proportions of responses to the 22 items by the sexes are presented in Table II.

Table II

Z Scores Describing the Differences in Proportions of Sick Responses to the 22 Items Between Receiving Ward and Plainwell Samples by Sex

Physical Disturbances

	Sex	
	Males	Females
1. Headaches, often	-1.281	.893
2. Trouble sleeping, often	.855	1.312
3. Hands tremble, often	1.924	1.330
4. Shortness of breath, often	1.849	.532
5. Cold sweats, often	1.478	.069
6. Heart beating hard, often	.204	1.330
7. Fainting spells, more than a few times	.114	1.065
8. Appetite, poor	.494	1.496
9. Weak all over, yes	1.967	2.652 *
10. Acid stomach, yes	2.132	1.312
11. Hot all over, yes	2.501 *	.107
12. Fullness in head, yes	.510	.540

Table II continued

	Sex	
	Males	Females
13. Worries get you down physically, yes	3.193 **	4.170 **
Non-Physical Disturbances		
1. Nervousness, often	1.906	2.154
2. Spirits, very low	1.543	2.658 *
3. Worrying type, yes	3.191 **	1.005
4. Feel apart, yes	2.690 *	2.092
5. Restlessness, yes	2.350 *	3.133 **
6. Memory all right, no	1.712	1.504
7. Couldn't get going, yes	3.346 **	.182
8. Nothing turns out, yes	2.668 *	3.718 **
9. Nothing worthwhile, yes	2.592 *	3.390 **

---

* These Z scores are significant at the .01 level for a one tailed test.

** These Z scores are significant at the .001 level for a one tailed test.

The results regarding the validity of the 22 items for the sexes are more complex than those for age. The Z scores in Table II tend to be largest for females. Of the 13 items pertaining to physical disturbances, eight of the Z scores are larger for the females, while

five of the nine Z scores in the non-physical disturbance category are largest for this sex. Most of the Z scores are not significant at the .01 level, however. When the significant Z scores are examined, the reverse of the above noted trend becomes apparent. Eight Z scores are significant for the male group, while six items significantly discriminate between the known groups for the females. Both sexes have only two physical items which prove to be valid, but six non-physical items discriminate for the males, while only four discriminate for the females.

In Chapter I, it was noted that although women tend to report physical symptoms more often than men, these symptoms need not be indicative of mental illness.[11] In the present study, it was found that the hospital and Plainwell males tend to report physical disturbances more often than females.[12] This finding contradicts those of other studies.[13] Moreover, the difference in findings cannot be readily attributed to differences in the nature of the symptoms,

---

[11] See Footnote 6 of Chapter I.

[12] See Appendix B, section IV.

[13] Commission on Chronic Illness, op. cit., p. 50, and Gerald Gurin, Joseph Veroff and Shiela Feld, Americans View Their Mental Health, New York: Basic Books, Inc., 1960, pp. 189-192.

since the items used in the studies are nearly identical.[14] We thus have a situation just the opposite of that which was assumed to exist.[15] In the present analysis, although the males tend to report more physical symptoms, most of these are not valid indicators of mental illness when examined by the known group method. Both sexes have only two out of a possible 13 physical items which discriminate between the mentally sick and well, demonstrating that the physical items are least apt to be indicative of mental illness regardless of sex or reporting.

More of the non-physical items are valid for the males than females. In a preliminary examination of the data, it was found that the males in the receiving wards of the hospital appear to be more mentally ill than the females.[16]

---

[14] The items of the 22 Item Mental Health Scale of the present analysis are almost identical in wording to those used by Gurin, et al., (Ibid., p. 184).

[15] An explanation of this finding is beyond the scope of this analysis due to the lack of relevant data. Apparently, females are not always more apt to admit physical symptoms. This appears to be true only for the hospital and Plainwell samples of the present analysis, however; the females in the Kalamazoo sample follow the expected trend. Perhaps these response differences are inherent in the nature of the samples themselves.

[16] In a preliminary analysis of the data, it was found that the receiving ward males had poorer mental health than the females when measured by the attendants' ratings. The chi-square is 9.25 which is significant at the .01 level for 2 degrees of freedom.

More items would be expected to discriminate for the males under these circumstances. It is interesting to note, however, that the sexes have the same number of valid physical items, in spite of the apparent poorer mental health of the males in the sick group. This finding lends further support to the suggestion that physical items are least apt to be valid indicators of mental illness. This is especially true for the males, in the present study.

The eight valid items for the males and the six valid items for the females constitute the Male Mental Health Scale and the Female Mental Health Scale, respectively.

Validity of the 22 Items by Educational Status

The proportions of positive responses to the 22 items were computed by educational status. The Z scores describing the differences between the proportions of the known groups are presented in Table III.

Table III

Z Scores Describing the Differences in Proportions of Sick Responses to the 22 Items Between Receiving Ward and Plainwell Samples by Educational Status

Table III continued

Physical Disturbances

	Educational Status	
	11 grades or less	12 grades or more
1. Headaches, often	-1.439	1.184
2. Trouble sleeping, often	.212	1.084
3. Hands tremble, often	1.250	1.642
4. Shortness of breath, often	.576	2.059
5. Cold sweats, often	.614	1.642
6. Heart beating hard, often	.362	.846
7. Fainting spells, more than a few times	.730	.846
8. Appetite, poor	.079	2.012
9. Weak all over, yes	.357	4.533 **
10. Acid stomach, yes	2.125	1.431
11. Hot all over, yes	.548	.895
12. Fullness in head, yes	-.811	1.431
13. Worries get you down physically, yes	3.028 *	5.035 **

Non-Physical Disturbances

1. Nervousness, often	1.024	3.000 *
2. Spirits, very low	1.858	2.341 *
3. Worrying type, yes	1.622	1.869

Table III continued

	Educational Status	
	11 grades or less	12 grades or more
4. Feel apart, yes	1.629	2.973 *
5. Restlessness, yes	2.790 *	2.496 *
6. Memory all right, no	2.461 *	1.442
7. Couldn't get going, yes	1.037	1.781
8. Nothing turns out, yes	2.571 *	3.399 **
9. Nothing worthwhile, yes	1.629	4.020 **

* These Z scores are significant at the .01 level for a one tailed test.

** These Z scores are significant at the .001 level for a one tailed test.

The Z scores in Table III clearly indicate that the items in the 22 Item Mental Health Scale are most discriminatory for the high educational status group. In the physical disturbance section, 12 of the 13 Z scores are largest for the high education group. The same trend is evident for the non-physical items. Here, seven of the nine Z scores are largest for the high educational status group. Only one of the physical items discriminates significantly for the low status group, while two prove to be valid for the high status group. Of the nine items in the non-physical section, three discriminate between the

known groups for the low education group, whereas six discriminate for the high educational status group.

These results demonstrate quite clearly that very few of the items in the 22 Item Mental Health Scale are capable of discriminating between the mentally ill and well samples for persons with low educational status. Twice as many items are valid for the high educational status group.

These findings appear to support the criticisms of the current definitions of mental illness. As noted in Chapter I,[17] many authors feel that the theoretical and operational definitions of mental illness contain middle class biases which tend to make the lower classes merely appear to be more mentally ill than the middle and upper classes. The results above do suggest that the items in the 22 Item Mental Health Scale may be treating the lower classes in an "unfair" manner.

The four valid items for the low educational status group and the eight valid items for the high educational status group make up the Low Educational Status Mental Health Scale and the High Educational Status Mental Health Scale, respectively.

---

[17] See Footnote 4 of Chapter I.

Summary and Conclusions

The results of the known groups item analysis support, for the most part, the first hypothesis of the study which states that the prevalence of symptoms, as measured by the items in the 22 Item Mental Health Scale, will vary by age, sex, and educational status, the physical items tending to be least valid for high age, female and low educational status groups.

It was found that, although more items discriminate between the known groups for the older persons, fewer items pertaining to physical disturbances are valid for this age group. It was suggested that this situation may introduce a bias into the 22 Item Mental Health Scale for this group.

The males have more valid items than the females, although the sexes have the same number of physical items which discriminate significantly between the mentally ill and well samples. The variations in item validity between the sexes occur within the non-physical symptom category. It was suggested that this may be due to the apparent difference in mental illness levels of the sexes in the hospital receiving wards.

The results pertaining to educational status and the validity of the 22 items show that the items are least discriminatory for the low educational status group, suggesting that the scale may be unfair

to the lower classes.

The over-all results show that most of the items in the 22 Item Mental Health Scale are of limited validity.[18] This is especially true for the items pertaining to physical symptoms.[19]

Six mental health sub-scales were constructed from the above procedures. Only those items which proved to be valid for a social sub-group were included in its scale.

---

[18] The results of the present analysis are to be compared with those of Manis, et al., (Jerome G. Manis, Milton J. Brawer, Chester L. Hunt and Leonard C. Kercher, "Validating a Mental Health Scale," American Sociological Review, Vol. 28, No. 1, February 1963, pp. 108-116. The results for this comparison are taken from the supplementary tables to the above article.) In the Manis study, 13 of the 22 items were found to discriminate between the receiving ward and Smallville samples at the .01 level for a one tailed test. These Z scores were computed for differences in proportions of sick responses between entire samples, rather than by age, sex and educational status, as in the present analysis. The more stringent controls of the present analysis result in fewer significant items. The nature of symptom variations by the sub-groups are additional fruits of the controlling procedures used in the present analysis.

[19] The interpretation of this finding must be qualified somewhat due to the policies of the staff at the Kalamazoo State Hospital. Patients who are physically ill are separated from the other patients and put in the "medical ward" of the hospital. This is done only if room is available, however, and patients with contagious diseases are considered first for such isolation. Whether or not such patients would report physical disturbances of the nature described by the 13 items of the 22 item scale more often than the other patients is an open question. Moreover, there is some evidence that physically ill respondents from the Plainwell sample have also been removed from the analysis. Of the eight refusals given by the Plainwell sample, five gave their reasons for not taking part in the interview. Of these five, two listed physical illness as the reason for their refusal.

Thus, whether or not the hospital policy reduces the discrimination between the samples on the 13 physical items is questionable. This is particulary true if one assumes that most of the physically ill in the Plainwell sample have also been removed. If this is the case, than the results pertaining to the physical items refer to comparisons between two physically well samples and should not be generalized to populations containing physically ill persons.

Chapter IV

VALIDATION OF THE MENTAL HEALTH SUB-SCALES

In the previous chapter, the validity of the items in the 22 Item Mental Health Scale was tested by age, sex and educational status sub-groups. Six mental health sub-scales were constructed from this procedure. The pragmatic validity of these sub-scales is the topic of the present chapter.

Three methods are used to test the validity of the sub-scales: a known group analysis between samples defined as mentally ill and mentally well, an independent criteria test within the mentally ill sample, and an independent criteria test within the mentally well sample. Attention is focused, first, upon the known groups analysis.

Validity of the Sub-scales by Known Groups Analysis

The known group technique used to test the pragmatic validity of the six mental health sub-scales is nearly identical to that used in the previous chapter. The major differences lie in the samples used and the nature of the measuring instruments.

In the present known group analysis, the hospital receiving ward patients are, again, defined as the mentally ill group, while the Kalamazoo respondents are classified as mentally well. As noted in

Chapter II, it is quite possible that some of the Kalamazoo respondents are mentally ill, or more ill than some, or all of the receiving ward patients. As a group, the Kalamazoo sample probably has better mental health, however.[1] The hospital receiving ward and the Kalamazoo samples are thus considered to represent the extremes of mental health, for the purposes of this analysis.[2]

Unlike the previous chapter, the present analysis focuses upon the validity of the mental health sub-scales as wholes. In other words, it is not the validity of the specific items which is the topic of this chapter, but rather the validity of the combinations of items to which the various sub-groups within the samples respond. Thus, although the known group procedures of the present chapter is fundamentally the same as that presented in Chapter III, the instruments with which it is concerned are not the same.[3]

---

[1] See Footnote 27 of Chapter II of this thesis.

[2] It is doubtful that these two samples actually represent the theoretical extremes of mental health. They represent the extremes only in the context of the present analysis.

[3] It may be argued that because the sub-scales are made up of the validated items, they are in part the same. This cannot be denied, but because all the scales contain more than one item, no scale can be equated with any one item, but rather represents a combination of items.

Each sub-scale is treated as a simple additive measure on which the respondents receive scores ranging from 0 to 8, depending on the number of items in the scale, and also on how the respondents answer the items. The scoring system is designed so that low scores represent that which is hypothesized to be good mental health, while higher scores designate progressively poorer mental health. Whether or not the scores on the sub-scales dc represent different levels of mental health is, of course, the topic of this chapter. If the sub-scales are valid indicators of mental illness, then persons who are known to be mentally ill should have higher scores than persons known to be mentally well. It is thus predicted that the receiving ward patients will have higher mental health scores on the appropriate sub-scales than the Kalamazoo respondents.

The mean mental health scores on the appropriate sub-scales for age, sex, and educational status sub-groups of the two samples are presented in Table IV.

Table IV

Mean Scores on the Various Mental Health Sub-Scales by Hospital Receiving Ward and Kalamazoo Samples

Table IV continued

Mental Health Sub-Scale	Receiving Ward Sample N	$\bar{X}$	Kalamazoo Sample N	$\bar{X}$
Low Age Scale	57	2.2	730	1.1
High Age Scale	46	2.8	453	1.3
Male Scale	55	3.4	284	1.6
Female Scale	50	2.3	899	1.0
Low Educational Status Scale	61	1.5	490	.8
High Educational Status Scale	37	3.3	678	1.1

The chi-square method was used to test the significance of the trends described by the mean scores on each sub-scale for each level of estimated mental health. The trends described by these means were transformed into 2 x 2 contingency tables. The scores on each sub-scale were dichotomized into high and low scores. The mental health criteria were also dichotomized. These procedures were used for Tables IV, V and VI of this chapter.

All the chi-squares in this table are significant well beyond the .001 level of confidence. See Appendix C for a listing of the chi-squares.

The mean scores in Table IV show very consistent relationships between the sub-scales and the known groups. The mean scores are larger for the receiving ward patients than the Kalamazoo respondents in every case.

Moreover, all the trends are significant well beyond the .001 level of confidence. These results clearly indicate that the sub-scales are quite capable of discriminating between the mentally ill and mentally well age, sex, and educational status sub-groups. Thus, not only are the individual items making up the scales valid, but the various combinations of the items also appear to be valid indicators of mental health. More will be said about these findings in the concluding section of this chapter.

Validity of the Sub-scales by Independent Criteria Analysis Within the Hospital Sample

Three independent estimates of mental health are used in the independent criteria test within the hospital sample. These estimates pertain to: the floor assignment of the patients by the hospital staff, which is based upon the severity of the mental illness of the patient, the attendant's ratings of the mental health of the patients under his supervision, and the number of previous admissions of the patient to the hospital for treatment. Each of these estimates was obtained independently of the patients' responses to the items making

up the sub-scales. Although these estimates are necessarily of a "face validity" nature,[4] they are defined as independent criteria of mental health in the present analysis.

The rationale of the independent criteria test is based upon the relationships between the criteria and the mean scores of the patients on the sub-scales. If the sub-scales do, in fact, measure mental health, then patients having less favorable mental health, as indicated by one or more of the independent criteria, should have higher scores on the sub-scales.

The mean scores on the appropriate sub-scales by age, sex and educational status sub-groups within the hospital for various levels of mental health, as measured by the independent criteria, are presented in Table V.

---

[4] This is perhaps the most basic paradox associated with validation studies. Seldom, if ever, are criteria, which are known to be valid, available with which to compare a measuring instrument. If such criteria were readily available, then there would be no need for the instrument. One must generally settle for approximate estimates. This sort of compromise with absolute definitions is not at all uncommon in scientific endeavers which deal with approximate truths. Although it is impossible to state that a given instrument is completely valid or invalid, one may judge the degree of validity of the instrument. This is generally accomplished by examining the "empirical connections" of the instrument in a wide variety of contexts. For an excellent discussion of this problem see Abraham Kaplan, The Conduct of Inquiry, San Francisco: Chandler Publishing Company, 1964, pp. 198-199. Also see William J. Goode and Paul K. Hatt, Methods in Social Research, New York: McGraw-Hill Book Company, Inc., 1952, pp. 238-239.

Table V

Mean Scores on the Various Sub-Scales by Hospital Receiving Ward Patients for Different Levels of Estimated Mental Health

Independent Estimates of Mental Health		Young		Old		Males		Females		Low Education		High Education	
		N	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$
Hospital Floor Assignment													
1	good	19	1.5	13	2.3	16	3.1	18	1.8	18	1.4	12	2.3
2	fair	23	2.4	22	2.9	25	3.2	20	2.7	28	1.5	16	3.4
3	poor	12	2.9	10	3.2	11	4.3	11	2.8	15	1.8	9	3.6
Attendant's Rating													
	good	24	2.3	11	2.2	15	2.4	21	2.2	19	1.1	14	3.1
	fair	23	1.9	20	2.8	21	3.1	22	2.4	25	1.5	17	3.6
	poor	9	3.0	13	3.1	17	4.6	5	2.6	16	2.1	6	3.0
Previous Admissions													
none	good	37	2.1	31	2.7	36	3.2	32	2.2	45	1.4	20	3.4
1 or more	poor	20	2.5	13	3.0	17	3.8	16	2.7	16	1.9	16	3.3

The chi-square method was used to test the significance of the trends described by the means in the table. No chi-square is significant at the .05 level of confidence. See Appendix C for the listing of the chi-squares.

Although not statistically significant, the trends described by the mean scores on the sub-scales are fairly consistent.[5] This is especially true for the floor assignment criterion of mental health. As the severity of mental illness as indicated by floor assignment, increases, the mean scores on the various sub-scales by the patients also increase.

The relationships between the mean mental health scores and the other two independent criteria are not as consistent as the foregoing. In the attendants' ratings section of the table, the young patients have their highest mean score in the "poor" mental health category as predicted. Their lowest mean score falls in the "fair" mental health category, destroying the ordered effect. Instead of the mean scores describing a progressive linear relationship as predicted, they form an inverse parabola.

The mean scores for the patients with high educational status, in the attendants' rating section, deviate even more from the expected pattern. Their highest mean score falls in the "fair"

---

[5] The reader is urged to distinguish between statistical significance and practical significance in the discussion of the trends in Table V. Nearly all of the trends run in the predicted fashion and this may be of practical significance, in and of itself.

category, while those having "good" mental health have higher mean scores than those having "poor" mental health.

Although most of the mean scores increase in magnitude as the number of previous admissions increases, the mean scores for the high educational status group; run in the opposite direction.

Most of the mean mental health scores on the various sub-scales describe trends which run in the predicted directions in this independent criteria analysis. There are deviations, however. It is possible that the discrepancies occuring within the attendants' rating section are due to the rather small number of patients falling into the "poor" mental health category, although the reversal of the order of the means in the previous admissions section for the high educational status group cannot be readily explained by a similar situation.

There is another factor which may reduce the relationships between the mental health scores and the independent criteria for the hospital patients. In a preliminary analysis of the data, it was found that the physical items of the 22 Item Mental Health Scale were able to discriminate between different levels of mental illness within the hospital receiving ward sample better than the

non-physical items.[6] It will be remembered that most of the physical items were excluded from the six mental health sub-scales, under consideration in this chapter, because they were not capable of discriminating between groups defined as mentally ill and mentally well.[7] This situation might explain why the sub-scales, being made up almost entirely of non-physical items, show less strength and consistency in their relationships with the various levels of mental illness within the hospital sample.

More will be said about this, after the results pertaining to the Kalamazoo sample have been presented.

Validity of the Sub-scales by Independent Criteria Analysis Within the Kalamazoo Sample

As a further test of the validity of the six mental health sub-scales, the mean scores on the scales were computed for age, sex,

---

[6] This preliminary analysis involved relating responses to the 22 items by hospital patients for different levels of mental health as estimated by an index constructed by this author. The index is based upon the interviewer's estimate of the patient's mental health, the attendant's rating of the patients' mental health and the floor assignment of the patient by the hospital staff. Of the 12 items which showed linear relationships with the index, nine were of the physical disturbance nature.

[7] These results were discussed in Chapter III of this thesis. It should be recalled that the known groups item analysis in Chapter III excludes all, or nearly all physically ill respondents.

and educational status sub-groups within the Kalamazoo sample by different levels of mental health which were estimated by two criteria. The data pertaining to these criteria were collected at the time of the interview.[8]

The first criterion of mental health refers to whether or not the respondent has ever seen a doctor about mental problems, or nervousness, while the second estimate involves a 45 item mental health scale which had been constructed by Manis, et al., in conjunction with consulting psychiatrists to the project.[9]

Inasmuch as the six mental health sub-scales reflect true differences in mental health levels, the mean scores on these scales should increase as the estimated mental health of the respondents becomes progressively poorer.

The mean scores on the various sub-scales by the Kalamazoo respondents by the various estimated mental health levels are presented in Table VI.

---

[8] This unavoidable situation casts some doubt upon the independence of these estimates. The interpretations of the results must therefore be made with caution.

[9] Jerome G. Manis, Milton J. Brawer, Chester L. Hunt, and Leonard C. Kercher, "Validating a Mental Health Scale," American Sociological Review, 28:1, February, 1963, p. 113.

Table VI

Mean Scores on Various Sub-Scales by Kalamazoo Respondents for Different Levels of Estimated Mental Health

Independent Estimates of Mental Health	Young		Old		Males		Females		Low Education		High Education	
	N	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$
Seen a doctor about mental problems?												
no good	593	.9	389	1.1	256	1.6	726	.9	396	.7	572	1.0
yes poor	137	1.8	64	2.1	28	2.2	173	1.6	94	1.2	106	1.9
45 Item Mental Health Scale *												
0	307	.5	218	.7	137	1.0	388	.5	180	.4	336	.5
1	129	.8	75	1.3	52	1.7	152	.8	93	.6	108	1.0
2	80	1.1	39	1.4	34	1.8	85	1.1	42	.9	77	1.4
3-6	136	1.6	79	1.9	41	2.6	174	1.4	102	1.0	111	1.9
7 and over	77	2.9	42	2.8	19	3.5	100	2.6	72	1.7	47	3.1

*The 45 Item Mental Health Scale is similar to the 6 sub-scales in that as the scores increase in magnitude, the mental health of the respondent is considered to become poorer.

The chi-square method was used to test the significance of the trends described by the means in the table. All the chi-squares are significant far beyond the .001 level of significance. See Appendix C for the listing of the chi-squares.

The mean mental health scores on the various sub-scales follow very strong and consistent patterns in Table VI. The lowest mean scores fall in the categories indicating the best mental health levels and become progressively larger as the estimated mental health of the sub-groups becomes poorer. All the trends run in the predicted directions, and all are significant far beyond the .001 level of confidence.

These results indicate the six sub-scales are indeed capable of discriminating between various levels of mental health within the mentally well sample, suggesting that the sub-scales are valid indicators of mental health.

Summary and Conclusions

The over-all results of the validation tests indicate the six sub-scales are valid measures of group mental health. Not all of the results are equal in strength and consistency, however.

The results for the known group analysis and the independent criteria analysis within the Kalamazoo sample are very strong and consistent with the predictions.

The trends in the independent criteria analysis within the hospital receiving ward sample are much weaker and less consistent than the foregoing. Although this situation may be due to the small

number of patients in some of the mental health categories, the nature of the sub-scales is probably an even more important factor.

Apparently, physical items are most discriminatory within the mentally ill group , while non-physical items prove most valid between the extremes of mental health, and also within the mentally well sample. There can be at least two explanations of this phenomenon: 1 - As a group approaches the very sick end of the mental health continuum, physical symptoms become more valid indicators of mental illness, and/or - 2 - the hospital staff tends to classify patients suffering from physical disturbances as mentally ill, when in fact, they are no more mentally ill than patients not exhibiting physical symptoms. To test these propositions is beyond the scope of this thesis, although they may serve as hypotheses for further research on the topic.

In any event, the validity of the sub-scales is qualified by the findings of this chapter. The sub-scales appear to be capable of discriminating between the samples defined as mentally ill and mentally well and also between different levels of estimated mental health within the mentally well sample. The validity of the sub-scales for the different levels of mental health within the mentally ill sample is much less definite.

Comparison of the validity of the six sub-scales, the 22 Item Mental Health Scale, and the 9 Item Mental Health Scale is the topic of the next chapter.

Chapter V

COMPARISON OF THE VALIDITY OF THE MENTAL HEALTH SUB-SCALES, THE 22 ITEM, AND THE 9 ITEM MENTAL HEALTH SCALES

In the previous chapters, a series of operations were performed which have led up to the topic of the present chapter. In Chapter III, six mental health sub-scales were constructed via the known groups item analysis for age, sex, and educational status sub-groups. The validity of these sub-scales was examined in Chapter IV. It was found that the sub-scales are quite capable of differentiating between samples defined as mentally ill and mentally well, and also between mental health levels within the mentally well sample. The relationships between the sub-scale scores and the independent criteria test within the mentally ill sample were less strong and consistent.

In the present chapter, the question as to whether or not the sub-scales are improvements over the 22 and 9 item scales is explored. More specifically, the validity of the sub-scales, the 22 item, and the 9 item mental health scales are compared by examining their strengths of relationship with several mental health criteria. Before discussing the results of the validity comparisons, a brief discussion of the rationale for these procedures is presented.

Rationale for Comparing the Validity of the Sub-Scales, the 22 Item, and 9 Item Mental Health Scales

Both the 22 and 9 item mental health scales were found to be reasonably valid measures of group mental health by Manis, et al.[1]

The procedures used to validate the six sub-scales in Chapter IV of this thesis are nearly identical to those used in the Manis study. The major difference in the procedures used in this and the Manis study is that the present analysis validated the sub-scales by age, sex, and educational status, while no such controls were introduced into the Manis study.[2]

It will be recalled that the rationale for controlling for age, sex, and educational status in the present analysis is based upon the possibility of the individual items of the 22 Item Mental Health Scale having biases associated with different age, sex, and educational characteristics. In other words, the items were thought to be contaminated by various characteristics other than mental health

---

[1] Jerome G. Manis, Milton J. Brawer, Chester L. Hunt and Leonard C. Kercher, "Validating A Mental Health Scale," American Sociological Review, 28:1, (Feb. 1963), pp. 108-116.

[2] It should be recalled that these controls were also used in the construction of the sub-scales in the present study.

specifically associated with different ages, sexes, and educational status levels. By controlling for these different social groups, it was hoped that at least some of these biases could be recognized and eliminated. The known groups item analysis, used in the sub-scale construction in Chapter III, did point out a number of response variations to the items by age, sex, and educational status. Most of the items in the 22 Item Mental Health Scale were not able to discriminate between the known groups to a significant degree for the six social sub-groups.

The basic problem posed in this thesis remains unanswered, however. If the six sub-scales have been, in fact, freed from biases associated with age, sex, and educational status, then they should prove to be more valid than the 22 item scale, which still includes the "biased" items. In other words, if the six sub-scales are more valid for the different social sub-groups, then they should be able to discriminate between various levels of mental health to a stronger degree than the 22 item scale.

There is an additional factor which should be taken into consideration at this time, however. Although a few items are common to one or more of the six sub-scales, variations between the items of the sub-scales do not appear to show any clear pattern, other than that most of the physical items of the 22 Item Mental Health Scale,

together with a few non-physical items, have been eliminated from all six sub-scales.[3] It is possible that, if the sub-scales do prove to be more valid than the 22 item scale, it is because they have merely been stripped of these commonly invalid items, and not because items, contaminated by behaviors specifically associated with each sub-group, have been excluded. In other words, there are a number of items, most of which are physical in nature, which have been excluded from all, or nearly all of the six sub-scales. These items may constitute the sole bias of the 22 item scale. If this is the case, then the construction of the six sub-scales would seem to have been an unnecessary task, since one scale, which has had the most consistently invalid items eliminated, from its protocol would be quite adequate.

The 9 Item Mental Health Scale provides a basis by which this possibility may be examined, since it contains, for the most part, only those items which have been shown to be most consistently valid for all six sub-scales.[4]

The question to be answered in the present chapter thus deals with the justification for having constructed the six sub-scales. If

---

[3] See Table I, II and III of Chapter III in this thesis.

[4] See Appendix A, sections V and VIII.

the sub-scales are more valid indicators of mental illness for the different age, sex, and educational status sub-groups, because biases associated with the characteristics of these sub-groups have been eliminated, then they should prove to be more valid than the 22 and 9 item mental health scales. The results are not completely consistent with this prediction, however.

Results of the Validity Comparisons

The validity criteria used in the previous chapter to validate the six sub-scales are used to make comparisons between the mental health scales in the present analysis.[5] In order to compare the validity of the various scales, a measure of strength of relationship is required. The Pearsonian C', or contingency coefficient was chosen for this purpose.[6]

The six sub-scales, the 22, and 9 item scales are related to each of the six mental health criteria for low and high age, male and female, and low and high educational status sub-groups.

---

[5] See pages 27-29 and 83-86 of this thesis for descriptions of these criteria.

[6] For a description of this statistic see Hubert M. Blalock, Social Statistics, New York: McGraw-Hill Book Company, Inc., 1960, pp. 230-231.

Contingency coefficients are computed for each of these relationships by which the validity comparisons are made.[7]

Validity comparisons between the mental health scales by known group analysis

The contingency coefficients describing the relationships between the mental health scales and the known groups criterion of mental health for each sub-group are presented in Table VII.

Table VII

Contingency Coefficients Describing Relationships Between the Mental Health Scales and Known Groups Criterion of Mental Health[a] For the Six Social Sub-Groups[a]

Sub-Group	Appropriate Sub-Scale	22 Item MH Scale	9 Item MH Scale
Low Age	.318	.234	.207
High Age	.338	.237	.337
Male	.355	.414	.409

[7] All of the coefficients are based upon 2 x 2 contingency tables and are thus comparable. There is, however, no test of significance for the differences between contingency coefficients currently available to put such comparisons on a more precise level. Until the mathematical statisticians create such a test, one must rely on mere inspection.

Table VII continued

Sub-Group	Appropriate Sub-Scale	22 Item MH Scale	9 Item MH Scale
Female	.288	.205	.213
Low Education	.308	.175[a]	.204
High Education	.441	.322	.322

[a]This coefficient is significant at the .01 level. All other coefficients are significant well beyond the .001 level. See Appendix C.

Comparison of the coefficients in Table VII shows that, in all but one case, the sub-scales are more strongly associated with the mental health criterion than are the 22 and 9 item scales. The coefficients between the criterion and the 9 item scale show the next most consistently strong relationships.

The Male Mental Health Sub-scale is the exception to this trend. The 22 Item Mental Health Scale shows the strongest relationship with the known groups criterion, the 9 item scale showing the next strongest, while the Male Mental Health Scale shows the weakest relationship for the male sub-group. Apparently, the 22 item scale is the most valid indicator for males, at least for this test.

This finding may be explained by a possible link between "masculine roles" and the 22 Item Mental Health Scale. The 22 item scale contains the greatest proportion of physical symptoms and it is possible that, inasmuch as males are less apt to admit physical weaknesses,[8] a male that does so may have a psychological problem. That is to say, he would be deviating from the "masculine norms," and this may itself be indicative of a potential, or actual mental problem.

The support for this interpretation is somewhat inconclusive in this study, however. It will be remembered that the Plainwell males reported physical symptoms more often than the females. The same was true for the hospital patients. The situation is different with the Kalamazoo sample, the males reporting physical disturbances less often than the females.[9] Moreover, the differences between the proportions of males reporting physical disturbances in the hospital

---

[8] For an excellent discussion on "compulsive masculinity" see Talcott Parsons, "Certain Primary Sources and Patterns of Aggression in the Structure of the Western World," pp. 274-284, in Patrick Mullahy, A Study of Interpersonal Relations, New York: Grove Press, Inc., 1949. Although Dr. Parsons is concerned primarily with aggression, he presents a description of the structures in our Western society which instigate and perpetuate "masculine psychology."

[9] See Appendix B, Section IV.

and Kalamazoo samples tend to be larger than those between the hospital and Plainwell samples.[10] Since the Male Mental Health Scale was constructed from items which discriminate between the hospital and Plainwell samples, few physical items were included. One might thus expect the 22 item scale to be more discriminating between the Kalamazoo and hospital samples, since it contains items which show large response differences between these samples, i.e., it contains the largest proportion of physical items.

In any event, the 22 Item Mental Health Scale is most discriminatory for the males, while the sub-scales for the other five social sub-groups prove to be the most valid in the present known groups analysis.

Validity comparisons between the mental health scales by independent criteria within the hospital sample

The coefficients describing the relationships between the mental health scales and the three independent criteria of mental health used within the hospital sample are presented in Table VIII.

---

[10] This is true for 8 out of the 13 physical items. See Appendix B, Section IV.

Table VIII

Contingency Coefficients Describing Relationships Between the Mental Health Scales and the Independent Criteria of Mental Health Within the Hospital Sample for the Six Social Sub-Groups.[a]

Sub-Group	Appropriate Sub-Scale	22 Item MH Scale	9 Item MH Scale
	Floor Assignment		
Low Age	.374	.444	.375
High Age	.203	.093	.068
Male	.046	.226	.183
Female	.337	.271	.218
Low Education	.087	.048	-.008
High Education	.440	.489	.451
	Attendants' Ratings		
Low Age	.005	.029	.082
High Age	.214	.370	.038
Male	.117	.403	.233
Female	.142	-.095	.002
Low Education	.134	.300	.155
High Education	-.186	-.048	.077
	Previous Admissions		
Low Age	.091	.098	.184

Table VIII continued

Sub-Group	Appropriate Sub-Scale	22 Item MH Scale	9 Item MH Scale
		Floor Assignment	
		Previous Admissions	
High Age	.114	.137	-.003
Male	.202	.039	.091
Female	.166	.157	.115
Low Education	.014	.064	.032
High Education	-.157	.036	.056

[a] None of the relationships are significant at the .01 level. See Appendix C.

The patterns of the contingency coefficients in Table III are not nearly as consistent as those for the known groups analysis, discussed above.

Three of the sub-scales show the strongest associations with the floor assignment of the patients in the hospital, while the 22 item scale shows the stronger relationships for the other three sub-groups, for this criterion of mental health. The 22 item scale has the

strongest association with the attendants' ratings for three of the six sub-groups, the 9 item scale showing strongest for two sub-groups. The sub-scales, the 22 item and the 9 item scales show stronger associations with the previous admissions criterion for two sub-groups each.

There is little pattern in the relationships between the mental health scales and the three mental health criteria in Table VIII, the relationships showing inconsistent trends across three criteria. This virtual lack of pattern may be due to the small numbers involved with the hospital sample. None of the coefficients are significantly related to any of the mental health criteria within the hospital sample, so that it is perhaps somewhat presumptuous to make comparisons between the mental health scales.

Keeping this in mind, the only trend which does appear in Table VIII with fair consistency is that the 22 Item Mental Health Scale tends to show the strongest, even if not significant, associations with the criteria for the greatest number of sub-groups. This finding is quite consistent with findings presented earlier, in which it was noted that physical items tend to be most discriminatory for different levels of estimated mental health within the hospital sample.[11]

---

[11] See Footnote 6 of Chapter IV of this thesis.

Since the 22 item scale contains 13 physical items, one might expect that this scale would prove to be most discriminatory within the hospital sample.

Validity comparisons between the mental health scales by independent criteria within the Kalamazoo sample

The patterns described by the coefficients in Table IX show some unexpected results. The sub-scales tend to show the weakest relationships with both criteria of mental health for the Kalamazoo Sample.

Table IX

Contingency Coefficients Describing Relationships Between the Mental Health Scale and the Independent Criteria of Mental Health Within the Kalamazoo Sample for the Six Social Sub-Groups[a]

Sub-Group	Appropriate Sub-Scale	22 Item MH Scale	9 Item MH Scale
	"Ever seen a doctor about mental problems?"		
Low Age	.364	.424	.456
High Age	.342	.347	.305
Male	.168[b]	.177[b]	.161[b]
Female	.279	.428	.470
Low Education	.234	.480	.458
High Education	.348	.264	.396

Table IX continued

Sub-Group	Appropriate Sub-Scale	22 Item MH Scale	9 Item MH Scale
		45 Item MH Scale	
Low Age	.453	.653	.644
High Age	.480	.480	.602
Male	.480	.420	.521
Female	.428	.578	.655
Low Education	.435	.597	.646
High Education	.521	.562	.604

[a] The chi-squares for these relationships are presented in Appendix C.

[b] These relationships are significant at the .01 level. All other relationships are significant well beyond the .001 level.

The 22 Item Mental Health Scale shows the strongest association with the item referring to whether or not the respondent has seen a doctor about mental problems for the high age, male and low educational status sub-groups. The 9 item scale, on the other hand, is most strongly related to this criterion for the low age, female and high educational status sub-groups.

Since the major difference between the 22 and 9 item scales is that the former includes a greater proportion of physical items, one might surmise that persons who are more apt to report physical disturbances, via the 22 item scale, are also more likely to have seen a doctor about mental problems. This, however, is not the case. Although the older and lower educated sub-groups of the Kalamazoo sample report more physical disturbances than the younger and better educated, they are less apt to have seen a doctor about mental problems.[12] The Kalamazoo males report fewer physical symptoms than the females and also report having seen a doctor for mental problems less often.[13] The relationship between physical symptoms and seeing a doctor for mental problems is thus not a simple one.

An explanation of the finding in Table IX may lie in the motivation of the person going to see the doctor in the first place. It may be that the older, male, and lower educated persons who went to see a doctor did so primarily for physical reasons, while the younger, female, and higher educated saw a doctor for more strictly psychological reasons. Thus, it may be that, except for males, those persons

---

[12] See Appendix B, Section IV and also Table VI of Chapter IV.

[13] See Appendix B, Section IV and also Table VI of Chapter IV.

who go to see a doctor for physical problems, and only incidentally psychological problems, are more apt to report physical disturbances to an interviewer, via the 22 item scale. The question as to why males do not follow this pattern may lie in their reluctance to admit physical disorder to an interviewer, although they may have seen a doctor about such ailments as well as for psychological reasons.

The foregoing statements are mere speculation and cannot be tested with the data at hand. Moreover, the above described situation, insofar as it exists, would not explain why the 9 item scale shows stronger associations with the "seeing a doctor" item than the sub-scales, since both types of scales have very few physical items.

The 9 item scale is even more consistent in its stronger relationships with the 45 Item Mental Health Scale. The only exception to this trend is for the younger sub-group which shows a slightly stronger relationship for the 22 item scale and the 45 item scale.

The over-all results in Table IX thus point out that the 9 Item Mental Health Scale appears to be the most valid indicator of mental illness within the Kalamazoo sample. That the 9 item scale tends to be more discriminating than the 22 item scale, with the exceptions noted above, would seem quite understandable, given the arguments

against the validity of items based upon physical disturbances.[14] The reasons for the 9 item showing showing superior discriminatory ability to the sub-scales is not so readily explainable. This is particularly true in view of the results of the known group analysis in which the sub-scales proved to be the most consistently valid.

Scale Validity, the Samples, and Validity Criteria

Apparently, whether or not one scale proves to be more valid than another depends a great deal upon the samples and criteria used in the validation procedures. The six sub-scales tend to be most valid, as indicated by their ability to discriminate between samples defined as mentally ill and mentally well; the 22 Item Mental Health Scale tends to show stronger, although not significant, relationships with the hospital independent criteria, while the 9 Item Mental Health Scale shows the strongest relationships with the criteria used in the analysis within the mentally well, or Kalamazoo sample.

The reasons for these different validity results may be based upon at least two general factors. These are the nature of the

---

[14] See Louisa P. Howe, "Problems in the Evaluation of Mental Health Programs," in R. Kotinsky and H. Witmer, Community Programs for Mental Health, Cambridge, Massachusetts: Harvard University Press, 1955, p. 287.

samples, and the mental health criteria used in the analysis.

The samples

The hospital receiving ward and the Kalamazoo samples vary on several characteristics. Moreover, the Plainwell sample, used in the sub-scale construction procedures, differs from both the hospital and Kalamazoo samples in some respects. Perhaps the most obvious difference between the samples is that the receiving ward patients are hospitalized presumably for mental illness, whereas the Kalamazoo and Plainwell samples are not. This was used as the basis of the definition of the mentally ill and well in the known groups analyses. There are other differences between the hospitalized and non-hospitalized samples, however. Although the Plainwell and hospital samples are quite comparable in their age distributions, the Kalamazoo sample is a relatively young sample. The hospital sample, on the other hand, has a nearly equal number of males and females, while the sex ratios for the Plainwell and Kalamazoo samples are somewhat lower. Moreover, the hospital sample has a much greater proportion of persons in the low educational status category than do the Plainwell and Kalamazoo samples.[15] Since sex is the only variable of

[15] See Appendix B, Sections I, II and III.

these three which can be completely controlled by dichotomizing, a basic weakness remains in this study which may have ramifications all the way back to the construction of the sub-scales. Yet to control for all possible levels of age and educational status was impossible in the present study.

The responses to the items of the 22 Item Mental Health Scale also vary by each sample.[16] One of the more outstanding response variations pertains to the reporting of physical items by sex. The males, it will be recalled, tend to report physical symptoms in the Plainwell and hospital samples more than the females, whereas the reverse situation is true for the Kalamazoo sample. The possible implications of this situation were discussed in an earlier section.

Of course, these comments do not begin to exhaust the possible number of variations between the samples. Possible rural-urban differences, for example, have not even been touched upon because of insufficient data. There are undoubtedly other factors which vary by sample, some of which may be influencing the results, not only of this chapter, but also those of previous chapters.[17] An examination

---

[16] See Appendix B, Section IV.

[17] Factors such as race, religion and ethnic background have not been considered in this thesis, for example. Nor has the difference between the relatively regimented life in an institution as compared

of all these possibilities as far beyond the scope of this, or any other single study.

The criteria

The second major factor which may have influenced the results of the present chapter pertains to the nature of the validity criteria.

The known groups analysis used to construct and then validate the six mental health sub-scales is based upon the assumption that the hospital receiving ward patients are more mentally ill, as a group, than the Plainwell and Kalamazoo samples. Although some evidence was presented in Chapter II supporting this assumption,[18] the definition of mental illness in the known groups analyses is ultimately based upon the hospitalization status of the samples. Furthermore, the known groups analyses may be yielding similar results because the procedures are very similar. Although the testing of the discriminatory ability of individual items is logically different from testing combinations of such items in a scale, one can argue that since each validated item is included in one or more

---

to the less structured routines of living outside an institution, such as the mental hospital, been considered. Yet, these and other factors could account for some of the response variations of the samples.

[18] See pages 23 and 26 and also footnote 27 of Chapter II.

scales, it is, in part, identical to the scales. This inherent similarity of the two procedures may thus serve to produce similar results for the two analyses.

Perhaps an even more important factor related to the known groups analyses, which may be effecting the results, is that both analyses use the hospital receiving ward patients as the mentally ill group. The two analyses are then not really independent of each other, since they share a common sample. Although the known groups analysis of the present chapter provides a reasonably logical test of the validity of combinations of items (i.e., the sub-scales), it is not quite independent of the sub-scale construction procedures.

The independent criteria within the hospital sample are based upon entirely different factors. The floor assignment criterion is based upon the hospital staff's judgments as to the severity of the mental illness of a patient. The attendants' ratings are of similar quality, being based upon their presumed acquaintance with the patients' mental conditions. The previous admissions criterion refers to the number of times the patient has come, or been sent to a hospital for treatment. All of these criteria are based upon "experts" opinions of the severity of the mental illness of the patients. Whether or not such decisions by these judges are contaminated by

their medical orientation was discussed briefly in Chapter IV.[19] It was noted that the physical items of the 22 item scale discriminated most effectively between the various estimated levels of mental illness within the hospital sample. A demonstration of this possible medical bias on the part of the hospital staff is beyond the scope of this analysis, however. It may be that physical items are, in fact, more indicative of mental illness within hospitalized groups.

The independent criteria used within the Kalamazoo sample are also subject to some criticism. These estimates of the mental health of the Kalamazoo respondents are perhaps the least independent of the various mental health scales tested for validity in the present chapter, since the data for the criteria and the scales were collected at the same time by the same interviewer.

The item inquiring as to the respondents' having seen a doctor about mental problems is perhaps most similar to the previous admissions criterion of the hospital patients. Both involve a contact with a professional who is presumably capable of alleviating the person's suffering. They differ in that the "seeing a doctor" item is a second hand report through an interviewer, whereas an

---

[19] See page 63 of this thesis.

admission to the hospital is an objectively recorded fact which does not rely upon one's memory or willingness to report such an occurrence.

The 45 item scale is still another type of mental health estimate. The nature of this scale appears to be most similar to the 9 item scale and the six sub-scales of the present analysis.[20] Thus, not only were the data for this scale collected at the same time as the data for the scales examined in this analysis, but it is also very similar in nature to all but the 22 item scale. It is quite possible that a respondent may balk at responding positively to an item which inquires whether or not he has seen a doctor about mental problems, since this might seem to come "dangerously" close to admitting that he is "going nuts" or what ever term such a person might use. Yet this same person may readily admit, or report a number of symptoms without the same fear, inasmuch as he does not recognize the symptoms as being indicative of mental illness. If this is the case, then the 45 Item Mental Health Scale may be merely a reliability check on the six sub-scales and the 9 item scale. That is to say, it may be merely indicating the degree to which a person is apt to admit similar symptoms.

---

[20] See Appendix A, section V and VIII.

Conclusions

The samples and mental health criteria used to test the validity of the various mental health scales are thus somewhat different from each other. Moreover, the validity criteria are of questionable validity themselves. As noted earlier, this is the most basic problem associated with all validity studies and is apparently insurmountable at the present time.[21] The investigator is thus forced to choose among the criteria and samples available, perform his operations, and then cautiously state his conclusions.

The variations in the strengths of validity between the various mental health scales may result from a number of factors. Some of these factors may stem from the limitations of the samples and criteria listed above. The sub-scales may relate most strongly to the known groups criterion of mental illness due to the similarity between this analysis and that used to construct the six sub-scales. The 22 item scale may show the strongest associations with the hospital criteria because of a medical bias on the part of the hospital staff. The 9 item scale may be most strongly related to the criteria used

---

[21] See Footnote 4 of Chapter IV of this thesis.

within the Kalamazoo sample because the lack of independence of these criteria from, and also their similarity to, the 9 item scale.

On the other hand, it may be that different scales are needed to measure the mental health of different groups. If one is interested in comparing the mental health of hospitalized and non-hospitalized groups, or wishes to predict which non-hospitalized groups are likely to be hospitalized for mental illness, then the sub-scales, constructed in the present study, might be most profitably used. The major exception to this generalization is the Male Mental Health Scale. The 22 Item Mental Health Scale appears to be the most valid indicator for males, and also for patients in the mental hospital. The 9 Item Mental Health Scale, on the other hand, seems to be most discriminatory between different levels of mental health within the mentally well sample.

There is still at least one more possible interpretation of the results of the present chapter. Although all the six criteria of mental health have been shown to have some limitations, some may be more handicapped than others. If one assumes that mental health is a continuous variable, then a criterion which involves the entire continuum would seem to be less limited than one which covers only a small segment of the characteristic. That is to say, a criterion which deals with one extreme, or the other, of the mental health

continuum would seem to be more limited than one which considers both extremes. Using this logic, the known groups criterion of mental illness appears to be the least limited, since it utilizes both extremes of mental health, as defined in this analysis.[22]

When the results of the present chapter are examined from this point of view, the sub-scales appear to be the best over-all indicators of mental health. The exception to this conclusion is the Male Mental Health Scale. As noted above, the 22 item scale proves to be more valid for males. It was suggested earlier, that this variation by sex may stem from "masculine psychology" in which the admission of physical weaknesses may be particularly indicative of mental illness for men.[23]

Summary

Comparisons of the validity of the six sub-scales, the 22 Item Mental Health Scale and the 9 Item Mental Health Scale were made. These comparisons were accomplished by examining the strengths of

---

[22] It will be remembered that the hospital sample represents the mentally ill extreme, while the Plainwell and Kalamazoo samples represent the mentally well extremes of the mental health continuum in the known groups analyses of this study.

[23] See Footnote 8 of this chapter.

relationships between the various mental health scales and six different criteria of mental health.

It was found that the sub-scales tend to discriminate best in the known groups analysis, with the exception of the Male Mental Health Scale, the 22 item scale discriminates best for males of the known groups and also within the hospital sample, while the 9 item scale shows the strongest associations with the criteria used to estimate the mental health for sub-groups within the non-hospitalized, or Kalamazoo sample. The criteria used within the mentally well sample tend to show the strongest associations with the mental health scales, the known groups analysis criterion the next strongest, while the criteria used within the hospital sample were the most weakly related to the scales and showed no relationship significant at the .01 level.

Chapter VI

SUMMARY AND CONCLUSIONS

Purpose of the Study Re-stated

The topic considered in this thesis was the construction and validation of six mental health sub-scales which were based upon procedures controlling for age, sex, and educational status. Each item in the a 22 item mental health scale was tested for its validity by age, sex, and educational status. Six sub-scales, one for each dichotomy of the three social sub-groups, were constructed from this procedure. The validity of these sub-scales was then tested via six different criteria of mental health. Finally, the validity of the six sub-scales, the original 22 item scale, and a related 9 item scale were compared by examining the strengths of relationship between these scales and the six mental health criteria.

The 22 Item Mental Health Scale had been validated in a previous study by Manis, et al.[1] They found that the scale was a reasonably

---

[1] Jerome G. Manis, Milton J. Brawer, Chester L. Hunt and Leonard G. Kercher, "Validating a Mental Health Scale," American Sociological Review, 28:1, 1963, pp. 108-116.

accurate measure of group mental health.[2] The validation procedures used in the Manis study did not include controlling for the possible effects of extraneous factors such as age, sex, and educational status, however. It was noted by the present author that an item in the 22 item scale might be tapping any number of factors other than mental health. Several examples were given of items which were thought to measure, at least in part, behaviors and/or conditions commonly associated with different age, sex, and educational status groups. The items classified as indicating physical disturbances of the respondents were given special attention in terms of their possible bias.[3]

Thus, although the 22 Item Mental Health Scale and its individual items were validated for entire samples, whether or not this scale and its items are equally indicative of mental illness for different age, sex, and educational status sub-groups making up the samples remained an unanswered question.

---

[2] The concept of group mental health should be distinguished from that which pertains to the mental health of individuals. Manis, et al., (Ibid.) used average scores on the 22 item and 9 item scales as indicators of the mental health level of groups. The present analysis used average scores and also proportions of persons falling between certain scale scores as indicators of group mental health.

[3] A number of references of criticisms of the various definitions of mental illness were given in Chapter I. See Footnotes 4 and 6 of Chapter I.

Hypotheses re-stated

Although the present analysis was somewhat exploratory in nature, two general hypotheses were presented. The first hypothesis deals with variations in symptoms by certain social sub-groups. This hypothesis was stated as follows:

I. It is predicted that the prevalence of symptoms, as measured by the items in the 22 Item Mental Health Scale, will vary by age, sex, and educational status, the items pertaining to physical disturbances proving least valid for the aged, female, and low educational status groups.

The second hypothesis pertains to the validity of the six sub-scales and is re-stated below.

II. The six mental health sub-scales, in which variations in validity by age, sex, and educational status have been controlled, will prove to be more valid than the 22 and 9 item scales.

Summary of Procedures and Findings

Known groups item analysis and sub-scale construction

In order to test the first hypothesis, a known groups item analysis was performed. Each item in the 22 item scale was tested for its ability to discriminate between samples defined as mentally

ill and mentally well.[4] Age, sex, and educational status were controlled for in these procedures.

The prevalence of symptoms, as measured by the items in the 22 item scale, did vary by age, sex, and educational status. Only a few items proved to be valid for more than one social sub-group.

Although more items discriminated between the known groups for the older respondents, fewer items pertaining to physical disturbances were valid for this age group. This part of the first hypothesis was thus supported by the findings.

The results pertaining to the sexes were not so consistent with hypothesis I, however. Although the males had more valid items than did the females, the sexes had the same number of physical items which proved to be capable of discriminating significantly between the known groups. Thus, although the prevalence of symptoms did vary by sex, the physical items did not prove to be less valid for the

---

[4] In the known groups item analysis, the Plainwell sample was defined as the mentally well group while the mental hospital receiving ward patients were considered to be mentally ill. This analysis should be distinguished from the known groups analysis used to validate the six sub-scales. The latter procedure used the Kalamazoo sample as the mentally well group.

females, as predicted.[5]

The results regarding the validity of the items by educational status showed that the items are least valid for the low educational status group. This was true for both the physical and non-physical items.

The over-all results of the known groups item analysis showed that most of the items of the 22 Item Mental Health Scale are of limited validity. This was found to be especially true for the items tapping physical disturbances of the respondents. The first hypothesis of the study was thus supported by nearly all the data in the known groups item analysis.

Six sub-scales were constructed from the above procedures for each dichotomy of the three social characteristics - age, sex, and educational status. Only those items which were found to be valid for a social sub-group were included in its scale.

Validation of the six mental health sub-scales

The validity of the sub-scales was ascertained by comparing the scores on the scales with six different estimates of mental health.

---

[5] It was noted that this unexpected finding may have stemmed from the fact that males in the Plainwell and hospital samples reported physical disturbances more often than did the females. This latter finding was also noted as being contradictory to previous findings of other studies. See pages 41 - 42 and also Footnotes 13, 14, and 15 of Chapter III.

The first estimate consisted of a known groups analysis very similar to the one used to test the validity of the individual items making up the scales.[6] The second, third and fourth estimates referred to the mental health of the receiving ward patients as indicated by the floor assignment, attendants' mental health ratings, and the number of previous admissions of the patients to the mental hospital. The fifth and sixth criteria pertained to the mental health levels of respondents in the mentally well sample. An item inquiring whether the respondent had ever seen a doctor about nervousness or mental problems constituted the first of these, while a 45 item mental health scale was used as the sixth criteria of mental health in the sub-scale validation procedures.

The over-all results of the sub-scale validation tests indicated that the sub-scales are valid measures of group mental health. Not all of the results were equal in strength or consistency, however.

The relationships between the sub-scale scores and the known groups criterion of mental health, and also those between the sub-scale scores and the two criteria used within the mentally well sample

---

[6] Care should be taken not to confuse this procedure with the sub-scale construction procedures. See pages 24-27 of this thesis.

were relatively strong and consistent. As the estimated mental health of the respondents, as measured by these criteria, became worse, their scores on the sub-scales indicated progressively poorer mental health. These relationships were all significant well beyond the .001 level of confidence.[7]

The results pertaining to the independent criteria of mental health used within the mentally ill sample were much weaker and less consistent than the foregoing. None of these relationships were significant at the .05 level of confidence. Fifteen of the eighteen relationships between the hospital criteria and the sub-scale scores did run in the predicted direction, however.

Comparison of the validity of the mental health sub-scales, the 22 item, and the 9 item mental health scales

Comparisons of the validity of the six sub-scales, the 22 item, and the 9 item scales were made for each of the six social sub-groups. These comparisons were accomplished by examining the strengths of relationship between the various mental health scales and the six

---

[7]Chi-square tests were used to test the significance of these trends.

different criteria of mental health, described in the section above.[8]

It will be recalled that the second hypothesis of the study predicted that the sub-scales would be more valid than the 22 and 9 item scales. That is to say, if the sub-scales were, in fact, freed from biases associated with different age, sex, and educational status groups, then they should be more strongly related to the six criteria of mental health. The results of the analysis were not completely consistent with this prediction, however.

It was found that the sub-scales tended to be most strongly related to the known groups criterion of mental health, while the 22 Item Mental Health Scale tended to show the strongest relationships with the criteria of mental health used within the hospital sample. The 9 Item Mental Health Scale, on the other hand, tended to show stronger associations with the criteria used within the mentally well sample than did the six sub-scales and the 22 item scale.[9]

These findings suggested that the strengths of the validity of the various mental health scales varied by the criteria used to determine

---

[8] Pearsonian C primes were used to describe the strengths of the relationships in the validity comparison procedures.

[9] Not all of the relationships followed these patterns. Most of the exceptions occurred within the hospital analysis, however, and no relationship in this portion of the analysis was significant at the .05 level. See tables VII, VIII, and IX and also pages 68-80 of Chapter V.

their validity. More will be said about this later.

Conclusions

Limitations of the Study

The conclusions of any study must necessarily be qualified by its limitations. The present analysis is no exception. A number of limitations of the data and procedures used in the study were mentioned throughout the thesis.

The nature of the samples was discussed at some length. It was noted that the samples varied on several characteristics other than mental health. The response rates to the 22 items by the sexes, for example, varied by sample, the Plainwell and hospital sample males reporting more physical disturbances than the females. The Kalamazoo respondents exhibited the reverse characteristics. The age, sex, and educational status composition of the three samples also varied, and since these variables were dichotomized, the effects of age and educational status could be only partially controlled.[10] Moreover, it was noted that the known groups item analysis dealt with samples from which all, or nearly all of the physically ill persons were removed.

---

[10] Sex is, by its very nature, a dichotomous variable so that its effects were completely controlled in the analysis.

How much error these limitations introduced into the results of the study remains unknown.

The procedures of the analysis also had some drawbacks stemming, not only from the foregoing, but also from characteristics inherent in their nature.

Two known groups procedures were used in the study. These procedures were basically different except that the samples defined as mentally well,[11] and the nature of the measuring instruments differed. The analysis used to construct the sub-scales dealt with the validity of the individual items of the 22 item scale, whereas the known groups analysis pertaining to the sub-scale validation and validity comparisons was concerned with the validity of the sub-scales as wholes. Since the sub-scales contain only items which proved to be valid in the item analysis, these procedures do have an inherent similarity and might thus be expected to yield similar results. Another limitation of the use of these two procedures in the same study is that both used the hospital receiving ward sample as the mentally ill group. The known groups analyses were thus not completely independent of each other.

The criteria used within the mentally ill sample were themselves of questionable validity. All were based upon "experts" opinions of

---

[11] See Footnote 4 above.

the mental health of the patients, and the question as to a possible medical bias on the part of the judges was raised.

Finally, the criteria used within the mentally well sample were perhaps the least independent of the mental health scales since all the data were collected at the same time by the same interviewer. It was noted that these criteria may have been serving as reliability checks as much as tests of validity.

Concluding remarks

Seldom does an investigator have perfectly valid criteria with which to compare a measuring instrument. He must generally settle for approximate estimates. If the instrument should prove to relate to a large number of such criteria in a wide variety of contexts, then the investigator would be reasonably safe in assuming that he had a valid measure.[12]

The sub-scales constructed in this analysis related to all six criteria of mental health in the predicted manner for nearly every social sub-group thus satisfying the requirements of a valid instrument noted above. The strengths of the validity of the various mental

---

[12] Abraham Kaplan, The Conduct of Inquiry, San Francisco: Chandler Publishing Company, 1964, pp. 198-199.

health scales varied by the nature of the criteria used, however. The sub-scales proved to be most valid in the known groups analysis, the 22 item scale the most valid within the mentally ill sample, while the 9 item scale tended to be most strongly related to the criteria of mental health used within the mentally well sample. These findings can be interpreted in a number of ways.

It may be that the sub-scales did best in the known groups analysis because this procedure was so similar to the construction procedures. The 22 item scale may have been most discriminatory within the mentally ill sample due to a medical bias stemming from the judges' training.[13] The 9 item scale may have related most strongly to the criteria within the mentally well sample because of the lack of independence of these criteria from the mental health data.[14]

---

[13] It will be recalled that the 22 item scale has the largest proportion of physical items in its protocol and might thus be expected to relate positively to the hospital staff's judgments, inasmuch as these judgments are biased by the medical orientation of the staff's training.

[14] It was also suggested that the 9 Item Mental Health Scale was very similar to the 45 Item Mental Health Scale and might thus be expected to relate to this criteria more strongly than the 22 item scale. The question as to why the 9 item scale should be more strongly related to this criteria than the sub-scales remained essentially unanswered, however. See pages 78-80 and 85-86 of this thesis.

On the other hand, it may be that different scales are needed to estimate the mental health of different types of groups. That is to say, perhaps the nature of symptoms which differentiate between the extremes of mental health are different than those which discriminate between different levels of mental health within the extremes.[15]

One might also argue that, although all the validity criteria of the present analysis were shown to have some limitations, some may have been more handicapped than others. If one thinks of mental health as lying along a continuum ranging from very poor to very good mental health, then the criteria used within the extremes of this continuum might be considered the most limited since they do not begin to cover the entire range of mental health. Using this reasoning, the criteria used within the mentally ill and mentally well samples of the present analysis would be considered more handicapped than the known groups analysis which came closer to dealing with the entire range of mental health, as defined in this analysis.

When the results of the study are examined from this point of view, the sub-scales appear to be the best over-all indicators of

---

[15] The samples defined as mentally ill and mentally well were considered to represent the extremes of mental health for the purposes of this analysis. It is very doubtful that these samples actually represent the theoretical extremes of this characteristic, however. These samples thus represented extremes only within the context of this study.

mental health. These sub-scales, with the exception of the Male Mental Health Scale,[16] were most strongly associated with the known groups criterion of mental health. That is to say, they tended to discriminate best between samples approaching the extremes of the mental health continuum. Moreover, as noted above, they were quite consistently related to the other mental health criteria showing that they maintain their validity in more than one context.

As in the case of virtially all research, the findings of the present analysis need further testing. The validity of the six sub-scales should be re-tested using other samples and other criteria of mental health in order to remove the cloud of doubt which still surrounds them. Will their validity be maintained within an even wider variety of contexts? Will such sub-scales prove to be fruitful tools

---

[16] The Male Mental Health Scale was found to relate more weakly to the known groups criterion of mental health than either the 22 or 9 item scales. It was noted that males who admit physical weaknesses are more apt to be psychologically maladjusted given the nature of "masculine psychology" which frowns upon such behavior. The 22 item scale might thus be expected to be more valid for males under these circumstances. The 9 item scale also has more physical items than the Male Mental Health Scale and may relate more strongly to the known groups criterion for this reason. It is quite possible then, that a mental health scale consisting of a large number of items pertaining to physical disturbances is the best measure of mental illness for males.

in the study of the prevalence of mental illness? Such questions can be answered only by further inquiry. The author hopes that this study will provide a stimulus as well as a rough guideline for further research on the epidemiology of mental illness.

APPENDICES

Appendix A

DESCRIPTIONS OF THE ITEMS USED IN THE STUDY

I. Age: The age of each respondent was obtained at the time of the interview for all three samples. The ages were dichotomized into: 44 years or less, and 45 years or more.

II. Sex: The sex of each respondent was recorded at the time of the interview.

III. Educational Status: The educational status of each respondent was determined by his last grade completed in school. The completed grades were dichotomized into: 11 years or less, and 12 years or more.

IV. 22 Item Mental Health Scale:

	Item	Responses
* 1.	Are you ever troubled with headaches?	Often * Sometimes Never
* 2.	Do you ever have any trouble in getting to sleep or staying asleep?	Often * Sometimes Never
* 3.	Do your hands ever tremble enough to bother you?	Often * Sometimes Never
* 4.	Have you ever been bothered by shortness of breath when you were not exercising or working hard?	Often * Sometimes Never
* 5.	Have you ever been bothered by "cold sweats"?	Often * Sometimes Never
* 6.	Have you ever been bothered by your heart beating hard?	Often * Sometimes Never

22 Item Mental Health Scale continued

	Item	Responses
7.	Are you ever bothered by nervousness (irritable, fidgety, tense)?	Often * Sometimes ____ Never ____
* 8.	Have you ever had any fainting spells?	Never ___ A few times ___ More than a few times *
* 9.	How would you describe your appetite?	Poor ___ * Fair ___ Good ___ Too Good ___
10.	In general, would you say that most of the time you are in	High spirits ____ Good spirits ____ Low spirits ____ Very low spirits *
11.	I am the worrying type.	Yes ____ * ____ No ____
12.	I feel somewhat apart even among friends.	Yes ____ * ____ No ____
*13.	I feel weak all over much of the time.	Yes ____ * ____ No ____
14.	I have periods of such restlessness that I cannot sit long in a chair.	Yes ____ * ____ No ____
*15.	I am bothered by acid (sour) stomach several times a week.	Yes ____ * ____ No ____
16.	My memory seems to be all right.	Yes ____ No ____ * ____
*17.	Every so often I suddenly feel hot all over.	Yes ____ * ____ No ____
18.	I have had periods of days, weeks, or months when I couldn't "get going."	Yes ____ * ____ No ____

22 Item Mental Health Scale continued

Item	Responses
*19. There seems to be a fullness (clogging) in my head or nose much of the time.	Yes * No
20. Nothing ever turns out for me the way I want it to.	Yes * No
*21. I have personal worries that get me down physically.	Yes * No
22. I sometimes can't help wondering whether anything is worthwhile anymore.	Yes * No

* An asterisk in front of an item number indicates that the item refers to a physical disturbance, while an asterisk in a response category indicates a positive, or sick response to the item.

V. Mental Health Scales Constructed from 22 Item Scale:

Manis 9 Item Mental Health Scale

1. Nervousness, often
2. Spirits, very low
3. Worrying type, yes
*4. Weak all over, yes
5. Restlessness, yes
*6. Acid stomach, yes
7. Nothing turns out right, yes
*8. Worries get you down physically, yes
9. Nothing worthwhile, yes

Low Age Mental Health Scale

*1. Trouble sleeping, often
*2. Weak all over, yes
*3. Worries get you down physically, yes
4. Restlessness, yes
5. Nothing turns out right, yes
6. Nothing worthwhile, yes

High Age Mental Health Scale

*1. Acid stomach, yes
*2. Worries get you down physically, yes
3. Spirits, very low
4. Worrying type, yes
5. Feel apart, yes
6. Nothing turns out right, yes
7. Nothing worthwhile, yes

Male Mental Health Scale

*1. Hot all over, yes
*2. Worries get you down physically, yes
3. Worrying type, yes
4. Feel apart, yes
5. Restlessness, yes
6. Can't get going, yes
7. Nothing turns out right, yes
8. Nothing worthwhile, yes

Female Mental Health Scale

*1. Weak all over, yes
*2. Worries get you down physically, yes
3. Spirits, very low
4. Restlessness, yes
5. Nothing turns out right, yes
6. Nothing worthwhile, yes

Low Educational Status Mental Health Scale

*1. Worries get you down physically, yes
2. Restlessness, yes
3. Memory all right, no
4. Nothing turns out, yes

High Educational Status Mental Health Scale

*1. Weak all over, yes
*2. Worries get you down physically, yes
3. Nervousness, yes
4. Spirits, very low
5. Feel apart, yes
6. Restlessness, yes
7. Nothing turns out right, yes
8. Nothing worthwhile, yes

* These items refer to physical disturbances.

VI. Independent Criteria of Mental Health Within the Kalamazoo State Mental Hospital Sample

Floor assignment of the patients:

Each patient in the Kalamazoo State Mental Hospital is assigned to a floor according to the severity of his mental illness. The most severe cases are assigned to the 3rd floor, the next most severe to the 2nd floor, while the least severe mental cases are assigned to the 1st floor.

Attendants' mental health ratings of patients under their supervision:

Each attendant in the Kalamazoo State Mental Hospital was asked to rate each of the patients interviewed and under his supervision on a five point scale ranging from very good to very poor mental health.

Previous admissions:

The number of admissions to the hospital by each patient interviewed was collected from the hospital records.

VII. Independent Criteria of Mental Health Within the Kalamazoo County Sample

Seen a doctor item:

Each respondent in Kalamazoo County was asked, "Is there anyone who has been to see a doctor about mental problems or nervousness?" The data for the present analysis were recorded so that when a respondent said he had seen a doctor for mental problems he received a positive score, but when he answered that no one or some other member of his family had seen a doctor he was given a zero score. This variable is thus by its very nature dichotomous.

VIII. 45 Item Mental Health Scale:

Is there any person living in your home:

1. Who is always arguing and quarrelling with other people?
2. Who seems to have no interest in anything that's going on around him?
3. Who keeps complaining about different aches and pains?
4. Who is always pessimistic and unhappy?
5. Who seems restless, upset, and can't stay put?
6. Who has suddenly changed from what he used to be and now acts strangely and different?
7. Who has some peculiar habit that he often does over and over again?
8. Who often acts silly and queer?
9. Who seems to have no trust or affection for anyone?
10. Who sits and daydreams and doesn't seem to understand what is going on?
11. Who often giggles and laughs to himself for no reason?
12. Who has had trouble because of heavy drinking?
13. Who often acts like he is going to hit someone but always holds himself back?
14. Who often gets excited and wound up for no reason at all?
15. Who seems to have temper tantrums like a child?
16. Who worries about dying even though in good health?
17. Who imagines that he is superior to everyone else and that he has some special power?
18. Who often imagines that he sees or hears things that no one else sees or hears?
19. Who talks very strangely so that no one can understand what he is talking about?
20. Who seems to be very shy and afraid of meeting other people?
21. Who complains about some part of his body not working although physically nothing seems to be wrong?
22. Who has a whole set of strange ideas?
23. Who is often depressed and feeling low for no good reason?
24. Who has a 'hot temper'?
25. Who claims he is no good and feels guilty about everything?
26. Who ever talks about committing suicide?
27. Who is sometimes very happy and excited and then becomes very depressed and unhappy?
28. Who can't stop worrying about everything?
29. Who feels that the whole world is against him?
30. Who doesn't seem to have any feeling about anyone or anything?

45 Item Mental Health Scale continued

31. Who is unreliable and can't ever really be trusted?
32. Who seems to be acting more and more mixed up and confused?
33. Who often finds it very hard to relax and take it easy?
34. Who is more afraid of something bad happening than he should be?
35. Who is overly worried that other people talk about him or watch him?
36. Who has nightmares?
37. Who is often nervous and irritable?
38. Who often has headaches?
39. Who complains about pressure or feeling a tight band around his head?
40. Who often has cold sweats or feels hot all over?
41. Whose hands often tremble for no reason?
42. Who often wonders whether anything is worthwhile any more?
43. Whose heart beats fast or skips a beat?
44. Who often just can't seem to get going?
45. Who often has trouble getting to sleep or who wakes during the night and can't fall asleep?

When a respondent reported himself as having one of these symptoms, he was given a score of one. His scores on the items were then added yielding a summary score on the scale.

Appendix B

RESPONSES TO THE ITEMS BY SAMPLE(S)

	Samples		
	Kalamazoo	Plainwell	Hospital
I. Age:	f	f	f
44 years or less	730	42	54
45 years or more	453	42	47
total	1183	84	101
II. Sex:			
males	284	29	48
females	899	55	53
total	1183	84	101
III. Educational Status:			
11 years or less	490	36	61
12 years or more	678	47	36
total	1168*	83*	97*

* These totals are smaller than those above because "don't know" responses were excluded.

IV. 22 Item Mental Health Scale:

Percentage of Sick Responses to 22 Items by Sample for Low Age Group

Item	Samples		
	Kalamazoo	Plainwell	Hospital
Physical:	N = 730	N = 42	N = 54
1. Headaches	11.8%	4.8%	5.7%
2. Sleeping	7.1	2.4	18.9
3. Hand tremble	2.2	0.0	3.8
4. Shortness of breath	3.8	2.4	15.1
5. Cold sweats	2.1	0.0	9.4
6. Heart beats hard	3.7	0.0	9.4
7. Fainting spells	3.2	2.4	1.9
8. Appetite	3.8	0.0	1.9
9. Weak all over	7.7	2.4	28.3
10. Acid stomach	12.1	11.9	20.8
11. Hot all over	13.0	11.9	18.9
12. Clogging in head	17.0	19.0	26.4
13. Worries get you down physically	16.3	4.8	47.2
Non-Physical:			
1. Nervousness	23.2	14.3	32.1
2. Spirits	10.3	0.0	5.7
3. Worrying type	46.2	52.4	62.3
4. Feel apart	18.2	26.2	47.2
5. Restlessness	32.9	19.0	52.8
6. Memory	8.2	2.4	11.3
7. Couldn't get going	46.4	33.3	43.4
8. Nothing turns out right	11.9	9.5	32.1
9. Nothing worthwhile	31.5	11.9	41.5

Percentage of Sick Responses to 22 Items by Sample for High Age Group

Item	Samples		
	Kalamazoo	Plainwell	Hospital
Physical:	N = 453	N = 42	N = 47
1. Headaches	10.2%	15.8%	13.6%
2. Sleeping	15.7	26.3	29.5
3. Hands tremble	4.3	5.3	20.4
4. Shortness of breath	6.8	7.9	11.4
5. Cold sweats	1.6	5.3	6.8
6. Heart beats hard	9.3	13.2	13.6
7. Fainting spells	3.9	0.0	4.5
8. Appetite	2.1	2.6	9.1
9. Weak all over	12.0	15.8	29.5
10. Acid stomach	20.0	10.5	34.1
11. Hot all over	22.7	28.9	38.6
12. Clogging in head	18.2	28.9	31.8
13. Worries get you down physically	15.7	10.5	45.5
Non-Physical:			
1. Nervousness	21.1	18.4	38.6
2. Spirits	11.7	0.0	15.9
3. Worrying type	38.9	34.2	65.9
4. Feel apart	17.4	13.2	38.6
5. Restlessness	24.5	26.3	50.0
6. Memory	11.5	10.5	27.3
7. Couldn't get going	37.3	21.1	43.2
8. Nothing turns out right	10.2	10.5	47.7
9. Nothing worthwhile	25.6	18.4	45.5

Percentage of Sick Responses to 22 Items by Sample for Male Group

Item	Samples		
	Kalamazoo	Plainwell	Hospital
Physical:	N = 284	N = 29	N = 48
1. Headaches	7.4%	17.2%	7.8%
2. Sleeping	7.0	17.2	25.5
3. Hands tremble	2.8	0.0	11.8
4. Shortness of breath	3.2	3.4	17.6
5. Cold sweats	1.0	3.4	13.7
6. Heart beats hard	2.8	10.3	11.8
7. Fainting spells	1.8	3.4	3.9
8. Appetite	2.5	3.4	5.9
9. Weak all over	7.4	10.3	29.4
10. Acid stomach	15.8	10.3	31.4
11. Hot all over	7.4	3.4	25.5
12. Clogging in head	20.4	24.1	29.4
13. Worries get you down physically	7.7	3.4	45.1
Non-Physical:			
1. Nervousness	14.1	13.8	33.3
2. Spirits	14.1	0.0	7.8
3. Worrying type	34.2	27.6	64.7
4. Feel apart	17.6	13.8	43.1
5. Restlessness	33.8	24.1	51.0
6. Memory	7.7	6.9	21.6
7. Couldn't get going	29.9	10.3	47.1
8. Nothing turns out right	10.2	6.9	33.3
9. Nothing worthwhile	23.2	10.3	37.2

Percentage of Sick Responses to 22 Items by Sample for Female Group

Item	Samples		
	Kalamazoo	Plainwell	Hospital
Physical:	N = 899	N = 55	N = 53
1. Headaches	12.2%	5.9%	10.9%
2. Sleeping	11.2	11.8	21.7
3. Hands tremble	3.0	3.9	10.9
4. Shortness of breath	5.5	5.9	8.7
5. Cold sweats	2.1	2.0	2.2
6. Heart beats hard	6.7	3.9	10.9
7. Fainting spells	3.9	0.0	2.2
8. Appetite	3.3	0.0	4.3
9. Weak all over	4.8	7.8	28.3
10. Acid stomach	14.6	11.8	21.7
11. Hot all over	19.4	29.4	30.4
12. Clogging in head	16.2	23.5	28.3
13. Worries get you down physically	18.5	9.8	47.8
Non-Physical:			
1. Nervousness	26.3	17.6	37.0
2. Spirits	9.8	4.0	13.0
3. Worrying type	46.3	52.9	63.0
4. Feel apart	18.0	23.5	43.5
5. Restlessness	28.4	21.6	52.2
6. Memory	10.0	5.9	15.2
7. Couldn't get going	47.1	37.3	39.1
8. Nothing turns out right	11.6	11.8	45.7
9. Nothing worthwhile	31.1	17.6	50.0

Percentage of Sick Response to 22 Items by Sample for Low Educational Status Group

Item	Sample: Kalamazoo	Plainwell	Hospital
Physical:	N = 490	N = 36	N = 61
1. Headaches	16.1%	18.2%	8.2%
2. Sleeping	13.5	24.2	26.2
3. Hands tremble	4.7	6.1	14.8
4. Shortness of breath	6.9	9.1	13.1
5. Cold sweats	3.5	6.1	9.8
6. Heart Beats hard	7.6	12.1	14.8
7. Fainting spells	4.5	0.0	1.6
8. Appetite	4.7	3.0	3.3
9. Weak all over	15.7	18.2	21.3
10. Acid stomach	20.6	9.1	27.9
11. Hot all over	21.0	24.2	29.5
12. Clogging in head	19.4	39.4	31.1
13. Worries get you down physically	18.8	12.1	42.6
Non-Physical:			
1. Nervousness	29.6	21.2	31.1
2. Spirits	9.2	0.0	9.8
3. Worrying type	46.3	48.4	65.6
4. Feel apart	22.7	24.2	41.0
5. Restlessness	34.1	24.2	54.1
6. Memory	9.6	0.0	16.4
7. Couldn't get going	48.8	33.3	44.3
8. Nothing turns out right	16.9	15.1	41.0
9. Nothing worthwhile	35.1	24.2	41.0

Percentage of Sick Responses to 22 Items by Sample for High Educational Status Group

Item	Sample: Kalamazoo	Plainwell	Hospital
Physical:	N = 678	N = 47	N = 36
1. Headaches	7.7%	4.3%	11.1%
2. Sleeping	8.0	6.4	19.4
3. Hands tremble	1.8	0.0	5.6
4. Shortness of breath	3.4	2.1	13.9
5. Cold sweats	.7	0.0	5.6
6. Heart beats hard	4.3	2.1	5.6
7. Fainting spells	2.5	2.1	5.6
8. Appetite	2.1	0.0	8.3
9. Weak all over	4.6	2.1	41.7
10. Acid stomach	10.9	12.8	25.0
11. Hot all over	13.1	17.0	25.0
12. Clogging in head	15.9	12.8	25.0
13. Worries get you down physically	14.2	4.3	52.8
Non-Physical:			
1. Nervousness	18.9	12.8	41.7
2. Spirits	12.1	0.0	11.1
3. Worrying type	41.6	40.4	61.1
4. Feel apart	14.6	17.0	47.2
5. Restlessness	27.0	21.3	47.2
6. Memory	9.1	10.6	22.2
7. Couldn't get going	38.6	23.4	41.7
8. Nothing turns out right	7.4	6.4	36.1
9. Nothing worthwhile	24.9	8.5	47.2

V. Mean Mental Health Scores on the Various Mental Health Scales by Social Sub-Group and Sample*

Sub-Group	Appropriate Sub-Scale	22 Item MH Scale	9 Item MH Scale
	Kalamazoo Sample		
Low Age	1.1	3.2	1.8
High Age	1.3	3.2	1.7
Male	1.6	2.6	1.5
Female	1.0	3.4	1.9
Low Education	.8	3.9	2.2
High Education	1.1	2.7	1.5
	Plainwell Sample		
Low Age	.5	2.3	1.3
High Age	1.0	3.2	1.6
Male	1.0	2.2	1.1
Female	.7	3.1	1.6
Low Education	.5	3.7	1.9
High Education	.7	2.2	1.1
	Hospital Sample		
Low Age	2.2	5.5	3.2
High Age	2.8	6.6	3.6
Male	3.4	6.2	4.1
Female	2.3	5.9	3.5
Low Education	1.5	6.0	3.4
High Education	3.3	6.1	3.7

* The N's for these mean scores are presented in Appendix B sections I, II and III.

VI. Frequencies for Independent Criteria of Mental Health* Within Hospital Sample

Floor Assignment:	f
1st floor	34
2nd floor	45
3rd floor	22
total	101

Attendants' Ratings:	
Good	36
Fair	43
Poor	22
total	101

Previous Admissions:	
None	68
1 or more	33
total	101

Within Kalamazoo Sample

Seen a doctor item:	
yes	201
no	982
total	1183

45 Item Scale:	
0-1 scores	730
2 or more	453
total	1183

*These data are broken down by social sub-groups in Chapter IV, Tables I and II.

Appendix C

STATISTICAL TESTS OF SIGNIFICANCE

Chi-squares describing the relationships between scores on the various mental health scales and the six criteria of mental health by social sub-group

Sub-Group	Appropriate Sub-Scale	22 Item MH Scale	9 Item MH Scale
	1. Known Groups (Kalamazoo vs. Hospital)		
Low Age	41.90	22.14	17.15
High Age	30.22	14.42	29.98
Male	22.74	31.80	30.88
Female	40.99	20.41	21.85
Low Education	27.45	8.58	11.61
High Education	77.07	39.00	39.00
	2. Floor Assignment of Patients (Hospital)		
Low Age	4.28	5.89	4.31
High Age	.96	.19	.11
Male	.06	1.42	.94
Female	3.00	1.87	1.19
Low Education	.23	.07	.00
High Education	3.98	5.03	4.19
	3. Attendants' Ratings of Patients (Hospital)		
Low Age	.00	.02	.20
High Age	1.07	3.31	.03
Male	.38	4.59	1.53
Female	.51	.23	.00
Low Education	.55	2.77	.74
High Education	.65	.04	.11

Statistical Tests of Significance continued

Sub-Group	Appropriate Sub-Scale	22 Item MH Scale	9 Item MH Scale
	4. Previous Admission of Patients (Hospital)		
Low Age	.24	.27	.98
High Age	.30	.42	.00
Male	1.15	.04	.23
Female	.68	.61	.33
Low Education	.00	.12	.03
High Education	.45	.02	.06
	5. Ever seen a doctor about mental problems? (Kalamazoo)		
Low Age	51.89	72.24	84.34
High Age	28.17	28.99	22.04
Male	4.11	4.52	3.72
Female	36.50	90.51	111.28
Low Education	13.75	63.85	56.84
High Education	43.58	25.07	57.61
	6. 45 Item Scale Score (Kalamazoo)		
Low Age	83.21	197.35	189.13
High Age	58.90	58.77	100.01
Male	36.92	27.51	44.18
Female	90.64	180.61	243.74
Low Education	51.03	107.43	127.85
High Education	106.59	127.49	150.70

All of the statistics for this thesis were computed via the IBM 1620 Computer located at The Western Michigan University Computer Center, Kalamazoo, Michigan. The programs for the computer were written by the author.

BIBLIOGRAPHY

SELECTED BIBLIOGRAPHY

BOOKS

Dunham, H. Warren, Sociological Theory & Mental Disorder, Detroit: Wayne State University Press, 1959.

Eaton, Joseph, and Robert J. Weil, Culture and Mental Disorders, Glencoe, Ill.: The Free Press, 1955.

Faris, Robert E. L., and H. Warren Dunham, Mental Disorders in Urban Areas, Chicago: University of Chicago Press, 1939.

Goldhamer, Herbert, and Andrew Marshall, Psychosis and Civilization, Glencoe, Ill.: The Free Press, 1953.

Guring, Gerald, Joseph Veroff, and Shelia Feld, Americans View Their Mental Health, New York: Basic Books, Inc., 1960.

Hoch, Paul H., and Joseph Zubin, Comparative Epidemiology of the Mental Disorders, New York: Grune and Stratton, 1961.

Jaco, E. Gartly, The Social Epidemiology of Mental Disorders, New York: Russel Sage Foundation, 1960.

Jahoda, Marie, Current Concepts of Positive Mental Health, New York: Basic Books, Inc., 1958.

Kotinsky, Ruth, and Helen Leland Witmer, Community Programs for Mental Health, Cambridge, Mass., Harvard University Press, 1955.

Goode, William J., and Paul K. Hatt, Methods in Social Research, New York: McGraw-Hill Book Company, Inc., 1952.

Leighton, Dorthea C., John S. Harding, David B. Macklin, Allister M. Macmillan, and Alexander H. Leighton, The Character of Danger, New York: Basic Books, Inc., 1963.

Menninger, Karl, The Vital Balance, New York: The Viking Press, 1963.

SELECTED BIBLIOGRAPHY

BOOKS continued

Pasamanick, Benjamin, Epidemiology of Mental Disorders, Washington, D. C., American Association for the Advancement of Science, 1959.

Plunkett, Richard J., and John E. Gordon, Epidemiology and Mental Illness, New York: Basic Books, Inc., 1960.

Rose, Arnold M., Mental Health and Mental Disorder, New York: W. W. Norton and Company, 1955.

Selltiz, Claire, Marie Jahoda, Morton Deutsch, and Stuart W. Cook, Research Methods in Social Relations, New York: Rinehart and Winston, 1962.

United States Department of Health, Education and Welfare, Acute Conditions, Incidence and Associated Disability, Public Health Service, Series 10, Number 1, 1963.

Weinberg, S. Kirson, Society and Personality Disorders, New York: Pretice-Hall, Inc., 1952.

Wepman, Joseph M., and Ralph W. Heine, Concepts of Personality, Chicago: Aldine Publishing Company, 1963.

ARTICLES

Bateman, J. Fremont, and H. Warren Dunham, "The State Mental Hospital as a Specialed Community Experience," American Journal of Psychiatry, 105: 1948, pp. 445-448.

Belknap, Ivan, and Hiram J. Friedsman, "Age and Sex Categories as Sociological Variables in the Mental Disorders of Later Maturity," American Sociological Review, 14: 1949, pp. 367-377.

SELECTED BIBLIOGRAPHY

ARTICLES continued

Bronfenbrenner, Urie, and Edward C. Devereaux, "Interdisciplinary Planning for Team Research on Constructive Community Behavior," Human Relations, 5: 1952, pp. 187-203.

Clausen, John A., and Marian Radke Yarrow, "Paths to the Mental Hospital," Journal of Social Issues, 11: 1955, pp. 25-32.

_____ "Social Science Research in the National Mental Health Program," American Sociological Review, 15: 1950, p. 408.

Caudill, William, and Bertram H. Roberts, "Pitfalls in the Organization of Interdisciplinary Research," Human Organization, 10: 1951, pp. 12-15.

Davis, Kingsley, "Mental Hygiene and Class Structure," Psychiatry, 1: 1938, 55-65.

Dysinger, Robert H., "Mental Health in the United State," Annals of the American Academy of Political and Social Science, 286: 1953, (entire issue).

Felix, R. H., and R. V. Bowers, "Mental Hygiene and Socio-environmental Factors," Milbank Memorial Fund Quarterly, 26: 1948, pp. 126-147.

Gursslin, Orville R., Raymond G. Hunt, and Jack L. Roach, "Social Class and the Mental Health Movement," Social Problems, 7: 1959-60, pp. 1-11.

Henry, Jules, "The Formal Structure of a Psychiatric Hospital," Psychiatry, 17: 1954, pp. 139-151.

Hollingshead, August B., and E. Kirby, "Social Mobility and Mental Illness," American Sociological Review, 19: 1954, pp. 577-584.

SELECTED BIBLIOGRAPHY

ARTICLES continued

_____ and Frederick C. Redlich, "Social Stratification and Psychiatric Disorders," American Sociological Review, 18: 1953, pp. 163-169.

Kaplan, Bert, Robert B. Reed, and Wyman Richardson, "A Comparison of the Incidence of Hospitalized and Non-hospitalized Cases of Psychosis in Two Communities," American Sociological Review, 21: 1956, pp. 472-479.

Lemkau, Paul, Christopher Tietze, and Marian Cooper, "A Survey of Statistical Studies in the Prevalence and Incidence of Mental Disorders in Sample Populations," Public Health Reports, 58: 1943, pp. 1909-1927.

Macmillan, Allister M., "The Health Opinion Survey: Technique for Estimating the Prevalence of Psychoneurotic and Related Types of Disorders in Communities," Psychological Reports, 3: 1957, p. 325.

Manis, Jerome G., Milton J. Brawer, Chester L. Hunt, and Leonard C. Kercher, "Estimating the Prevalence of Mental Illness," American Sociological Review, 29: 1964, pp. 84-89.

_____ "Validating a Mental Health Scale," American Sociological Review, 28: 1963, pp. 108-116.

Mishler, Elliot G., and Nance E. Waxler, "Decision Processes in Psychiatric Hospitalization: Patients Referred, Accepted, and Admitted to a Psychiatric Hospital," American Sociological Review, 28: 1963, pp. 576-587.

Parsons, Talcott, "Illness and the Role of the Physician: A Sociological Perspective," American Journal of Orthopsychiatry, 21: 1951, pp. 452-460.

Pearlin, Leonard I., and Morris Rosenberg, "Nurse-Patient Social Distance and the Structural Context of a Mental Hospital," American Sociological Review, 27: 1962, pp. 56-65.

SELECTED BIBLIOGRAPHY

ARTICLES continued

Scheff, Thomas J., "The Role of the Mentally Ill and the Dynamics of Mental Disorder: A Research Framework," Sociometry, 26: 1963, pp. 436-453.

Woodward, Julian L., "Changing Ideas on Mental Illness and Its Treatment," American Sociological Review, 16: 1951, pp. 443-454.