

THE RELIABILITY AND VALIDITY OF GENERAL HEALTH RATINGS

John E. Ware, Jr.

November 1976

P-5720

The Rand Paper Series

Papers are issued by The Rand Corporation as a service to its professional staff. Their purpose is to facilitate the exchange of ideas among those who share the author's research interests; Papers are not reports prepared in fulfillment of Rand's contracts or grants. Views expressed in a Paper are the author's own, and are not necessarily shared by Rand or its research sponsors.

The Rand Corporation
Santa Monica, California 90406

Abstract

This article is a much abridged summary of a lengthy technical report regarding studies of general health perceptions conducted for the National Center for Health Services Research. Eight scales pertaining to perceptions of Prior Health, Current Health, Health Outlook, Resistance-Susceptibility to Illness, Health Worry/Concern, and Sickness Orientation as well as Rejection of Sick Role and Attitude Toward Going to the Doctor were constructed from 32 items in a standardized survey instrument that can be self-administered. Tests of the reliability, validity, and stability of rating scores are described. It is recommended that the scales be used in studies requiring general health measures, i.e., those tapping both objective and subjective health constructs and negative and positive well-being. Suggestions for future research are offered.

Acknowledgments

The author gratefully acknowledges the contributions of others during the research and the preparation of this paper: W. Russell Wright for supervising data-gathering activities in all field tests; Ann G. Karmos, Mary K. Snyder, and Anita L. Stewart for their assistance with the statistical analyses; Mary Hruby and Leo Reeder of the UCLA Survey Research Center for their assistance with the Los Angeles County field test; Cathy A. Donald for her contributions to the review of literature; William H. Lohr (the Project Officer at DHEW) for his cooperation throughout the research; Allyson Davies-Avery for her many helpful contributions during the writing of this manuscript; and Robert H. Brook, Shawn A. Johnston, and the editors of *Health Services Research* for their helpful comments during editing.

Data gathering and analyses were performed pursuant to Contract No. HSM 110-72-299, National Center for Health Services Research, U.S. Department of Health, Education, and Welfare (DHEW), Washington, D.C. Additional analyses and the preparation of this paper were performed pursuant to the Health Insurance Study, a grant from DHEW.

This paper was prepared for presentation at the Health Status Index Conference held in Phoenix, Arizona, October, 1976 and sponsored by the National Center for Health Services Research, U.S. Department of Health, Education, and Welfare, and an edited version will appear in *Health Services Research*, Volume 11, Number 4, Winter, 1976.

THE RELIABILITY AND VALIDITY OF GENERAL HEALTH RATINGS

John E. Ware, Jr.

The Rand Corporation, Santa Monica, California

I. INTRODUCTION

Self ratings of general health are frequently used in studies of relationships among health variables, health and illness behaviors, and population health status.¹ Examples include the widely used single-item rating of health in terms of "excellent," "good," "fair," or "poor," and endorsement of statements of opinion, such as "I have been feeling bad lately." These measures are general in that they do not focus on a single component of health, such as physiologic, physical, mental, or social health. Presumably, they tap a general health concept underlying all components of health. Furthermore, they operationalize health as a personal assessment rather than as a phenomenon that can be directly observed, such as a day spent in bed due to poor health. Thus, general health ratings are designed to tap both the objective and subjective aspects of health as perceived by respondents. Finally, properly constructed general health rating scales reflect individual and group differences throughout the continuum from negative health states to positive well-being and, therefore, may contribute to a comprehensive definition of health.

These characteristics and advantages of general health ratings presume that the survey instruments on which they are based satisfy basic psychometric criteria. However, despite the frequent publication of studies including general health ratings,²⁻²⁵ little information is available for purposes of a psychometric evaluation. Thus, it is not clear how general health ratings should be scored and what, if anything, they measure.

Published reports indicate confusion among investigators regarding the level of enumeration²⁶ achieved with general health ratings. From the statistical methods used, it appears that the same or very similar general health rating items have been enumerated as nominal,^{4,8,9,11-14,16-19}

ordinal,²⁰ and interval^{3,5-7,10,16,21-23} data. With some exceptions,^{2,9,15,24,25} ratings have been scored from responses to a single questionnaire item in the absence of evidence that single-item rating scores are reliable. Finally, systematic studies of the validity of general health ratings could not be found. In the absence of reliability data, it is impossible to evaluate published findings that pertain to validity. Thus, there exists little basis for assigning meaning to rating scores.

At the instigation of the National Center for Health Services Research (NCHSR), and in order to further understanding of general health ratings, numerous psychometric studies have been conducted. A general health rating instrument (Form I of the Health Perceptions Questionnaire) was developed and fielded. Results of the field test, including scaling studies using traditional methods, indicated the need for questionnaire revisions to satisfy reliability and validity standards.²⁷ This paper is a much abridged summary of a lengthy technical report²⁸ regarding the studies conducted for the NCHSR, and presents major findings for the revised instrument, including construction of general health rating scales, reliability of single-item and scale scores, and validity of scale scores in relation to a wide range of health-related variables.

II. METHODS

Data-Gathering Methods and Population Characteristics

Data were gathered both with and without interviewer supervision using Form II of the Health Perceptions Questionnaire (HPQ), a standardized self-administered survey instrument containing 36 items structured as statements of opinion regarding health and sick role propensity (see Table 1). The HPQ required approximately eight minutes to administer, on the average. The entire interview schedule, including standardized survey instruments used to gather data regarding demographic and socioeconomic characteristics of the respondents and health-related information used in validity analyses, required less than an hour to administer, on the average. (For the entire interview schedule, see 29.)

Analyses performed thus far were based on data from five field tests involving approximately 2,000 adult respondents sampled from general populations between 1973 and 1975. In all field tests, mixed sampling designs were used to draw representative households and one or more adults from each household were selected.²⁸ A summary of respondent characteristics is presented in Table 2.

The rate of complete and usable interviews was approximately 37 percent for the two field tests that employed mailout/mailback methods. Interviewer supervision of questionnaire administration resulted in rates of usable returns ranging from 62 to 95 percent. Complete and usable returns were received from approximately 67 percent in the field test that employed interviewer dropoff and mailed return of questionnaires. The effects of data-gathering methods and partial returns on conclusions about health from these field tests are discussed in detail elsewhere.²⁸ Based on that discussion, it is reasonable to assume that data-gathering methods and differences in return rates did not influence the conclusions drawn from the methodological studies described in this paper.

Plan of Analysis

The specific goals of the research were to: (1) confirm the item groupings (eight health perception scales) in Form II of the HPQ hy-

pothesized from the field test of Form I;²⁷ (2) evaluate the items in terms of scaling criteria to achieve roughly normally distributed continuous scale scores; (3) assess the reliability of item and scale scores and the long-term stability of scale scores; and (4) evaluate the validity of scale scores. To guard against population-specific results, many analyses were independently replicated using data from four of the five field tests.

Scaling of Items. Any study of the validity of general health ratings based on questionnaire data can only be as good as the questionnaire items that are employed. Results will reflect both the soundness of theory underlying hypothesized relationships and the adequacy of the measures. Therefore, to achieve the best possible measures, extensive evaluation of health perceptions scales using psychometric criteria was the first step in this research.

Prior to scaling the HPQ items, they were grouped according to the eight specific perceptions they were hypothesized to measure (number of items shown in parentheses): Current Health (10); Prior Health (4); Health Outlook (6); Health Worry/Concern (4); Resistance-Susceptibility to Illness (4); Sickness Orientation (2); Rejection of Sick Role (4); and Attitude Toward Going to the Doctor (2). The hypothesized item groupings (shown in column 1 of Table 1) were based on the results of factor analytic studies of Form I of the HPQ (described in detail elsewhere²⁷).

To verify that items in each group had highest loadings on one and the same factor and no other high loadings, correlations among items were factor analyzed independently for each field test using the Principal Factor Method.³⁰ After evaluation of unrotated solutions, they were rotated according to the Tandem Criteria Method, which makes use of the correlations among items during two rotations of the initial solution.^{30,31}

Items in each hypothesized grouping that had one high loading (± 0.40 or greater) on one and the same rotated factor in each field test were retained and used to score that factor (by computing the simple algebraic sum of item scores). For further evaluation of the retained and discarded items, criteria of discriminant validity³² were applied in evaluating item-scale correlations. In

order to retain the item at this step, each item-scale correlation (corrected for overlap as suggested by Howard and Forehand³³) was required to be higher for the scale (perception) the item was hypothesized to measure than for all other scales that employed the same method of measurement.

Scoring of Items and Scales. Each HPQ item was accompanied by five response categories: Definitely True, Mostly True, Don't Know, Mostly False, and Definitely False. Items were scored so that a high number defined the health perception indicated by the scale name. For example, high scores for items in the Current Health scale indicated perceptions of good current health while low scores indicated poor current health; high scores for items in the Health Worry/Concern scale indicated presence of health worry/concern while low scores indicated absence of health worry/concern. A score for each of the eight health perception scales was computed for each respondent using the simple algebraic sum of scores for items that satisfied factor analytic and discriminant validity criteria.

Reliability and Stability of Scores. Test-retest reliability estimates were obtained for both item and scale scores during each of two field tests by computing product-moment correlations between scores obtained approximately six weeks apart from the same respondents (N=183). Internal-consistency reliability for scales was estimated independently in four field tests using Cronbach's³⁴ Alpha coefficient and data obtained from all respondents. It should be noted that test-retest coefficients are biased downward whenever trait changes occur between administrations.

The long-term stability of selected perceptions measured by HPQ scales (Current Health, Resistance-Susceptibility to Illness, Prior Health, and Health Worry/Concern) was estimated from product-moment correlations between alternate forms of HPQ scales administered to the same respondents (N=92) approximately two years apart.

Validity of Scale Scores. In the absence of other measures of general health perceptions against which to judge the validity of the HPQ scales, two methods of validation were used. The first method consisted of factor analytic studies of correlations among the scales. The second method consisted of studies of the relationships among the scales

and other variables that should exist if the scales measure what they are supposed to measure. Both methods relied heavily on theory regarding health constructs and, therefore, constituted strategies for investigating the construct validity of the scales. In order to be certain that the validity of the measures was being tested and that results could not be questioned on theoretical grounds, only those hypotheses based on well-accepted theory and relationships involving health variables that had been well documented in the empirical literature were tested.

Using the first method, correlations among the eight scales were factor analyzed to derive higher order factors. These factors were interpreted on the basis of the manifest content of scales in relation to their loadings on the factors. Patterns of loadings across factors were then interpreted to evaluate the construct validity of each scale score, i.e., the extent to which each scale is a measure of each higher order factor. For these purposes, a matrix of product-moment correlations among the eight health perception scales was computed for each field test. Higher order factors were extracted and rotated to orthogonal simple structure using the methods described previously for item scaling. Because the higher order factor structures were nearly identical across field tests, it was possible to summarize the proportion of reliable variance in each scale accounted for by each factor across four field tests without loss of information.

Based on the findings for Form I of the HPQ,²⁷ a major unrotated factor pertaining to perceptions of general health and additional unrotated factors pertaining to future health and/or sick role propensity were hypothesized to emerge from the higher order factor analyses. It was hypothesized that those HPQ scales primarily measuring past and present general health (Prior Health and Current Health) would correlate highly (± 0.40 or greater) with the same rotated health factor and would not correlate highly with the factor(s) pertaining to future health and/or sick role propensity.²⁷ It was also hypothesized that the remaining HPQ scales (Health Outlook, Health Worry/Concern, Resistance-Susceptibility to Illness, Sickness Orientation, Rejection of Sick Role, and Attitude Toward Going to the Doctor) would correlate highly with the rotated factor(s) defining health outlook and/or sick role propensity, and would not correlate highly

with the rotated present/prior health factor. In addition, it was hypothesized that Resistance-Susceptibility to Illness and, to a lesser degree, Health Worry/Concern would have noteworthy secondary loadings on the rotated present/prior health factor. The latter hypotheses were based on studies of interrelationships among HPQ Form I scales and are consistent with the notion that perceptions of resistance-susceptibility and health worry/concern are intervening variables between perceptions of past/present health *and* future health (Health Outlook and Sickness Orientation) and sick role propensity (Rejection of Sick Role and Attitude Toward Going to the Doctor).

In the case of the second method used to examine validity, relationships among the HPQ scale scores and 11 other measures of health, seven health and illness behaviors, and age were evaluated. Definitions of the validity variables studied and specific hypotheses regarding their associations with favorable perceptions of health (as defined by the six HPQ health scales) are presented in Table 3. In summary, it was hypothesized that favorable health perceptions would be positively associated with positive definitions of health (e.g., psychological well-being), negatively associated with negative definitions of health (e.g., bed days), negatively associated with illness behavior (e.g., number of doctor visits), and would not be associated with health behavior (e.g., check-ups). The latter hypothesis was based on previous findings.^{20,27} Relationships were estimated using product-moment correlations and a cross-sectional study design.

Hypotheses regarding relationships between the two HPQ scales pertaining to sick role propensity (Rejection of Sick Role and Attitude Toward Going to the Doctor) and the validity variables studied were different from those stated in Table 3. First, it was hypothesized that these two scales would not correlate with the other health variables studied (except the behaviorally defined variables) because sick role propensity is not health. Significant relationships were hypothesized between the two HPQ sick role propensity scales and behaviorally defined health variables studied (such as bed days) because such behaviors reflect both health and the decision to be or not to be sick (i.e., sick role propensity). For the same reason, it was hypothesized that the two sick role propensity scales would also correlate with the health and illness behaviors studied (e.g., number of doctor visits).

Since a high score for Attitude Toward Going to the Doctor and a low score for Rejection of Sick Role define sick role propensity, it was hypothesized that the two scales would be positively and negatively correlated, respectively, with the behaviors studied.

III. RESULTS

Construction of Scales

Thirty-two of the 36 HPQ Form II items satisfied factor analytic and discriminant validity criteria; the four items that did not were eliminated. Rotated and unrotated factor solutions for items and results of discriminant validity tests are presented and discussed in detail elsewhere.²⁸ All eight hypothesized item groupings clearly appeared as factors in the factor analyses in all four field tests. Only 12 errors were observed in 896 tests of the discriminant validity criteria, i.e., 99 percent of the tests were successful for items. Interpretation of the manifest content of items in each group that satisfied these criteria served as the basis for assigning tentative names to the eight HPQ scales.

The 32 items that satisfied factor analytic and discriminant validity criteria were used to construct eight scales containing from two to nine items each, as defined in Table 1. Six of the eight scales were balanced, i.e., contained both favorably and unfavorably worded items. Two of the scales (Sickness Orientation and Rejection of Sick Role) were constructed entirely from items worded in the same direction.

Distribution of Scale Scores

A summary of descriptive statistics (means and standard deviations) for scores computed from HPQ scales across four field tests is presented in Table 4. Frequency distributions for scores are presented elsewhere.²⁸ The goal of roughly normally distributed scale scores was nearly achieved. Keeping in mind that the lowest possible score is the number of items in the scale (k in Table 4) and that the highest possible score is five times that number, it is clear from Table 4 that each scale mean was close to the midpoint of the scale range. Further, standard deviations were most often approximately one-fifth to one-sixth of the scale range. Means tended to be lower for HPQ health scales in the East St. Louis field test, in which a larger proportion of the sample was disadvantaged (and apparently less healthy).

Reliability: Single-Item Scores

Findings regarding the test-retest reliability of HPQ scores, each computed from a single item, indicated two clear trends.²⁸ First, item scores tended to be less reliable in the East St. Louis field test. Scores computed for 18 of the 32 items were not sufficiently reliable for purposes of group comparisons in East St. Louis, according to the 0.50 standard suggested by Helmstadter.³⁵ Seven of the 32 coefficients did not meet this standard in Sangamon County. This trend may reflect population differences in reliability and/or stability of the health traits measured by these items. Second, one clear trend in differences in coefficients between item groupings was apparent; test-retest coefficients for items in the Rejection of Sick Role group tended to be lowest in both field tests.

Reliability: Scale Scores

Internal-consistency and test-retest reliability coefficients for HPQ scales in four field tests are summarized in Table 5. Almost without exception, scale scores were sufficiently reliable for purposes of group comparisons (i.e., 0.50 or greater). In some instances, the Current Health scale was sufficiently reliable for purposes of individual comparisons, i.e., internal-consistency reliability coefficients equaled or exceeded 0.90. Only two of the 32 internal-consistency coefficients were below 0.60. Median coefficients (across field tests) ranged from a low of 0.59 for Rejection of Sick Role and Sickness Orientation scales to a high of 0.91 for the Current Health scale.

Given that reliability coefficients define the proportion of true score variance, they can be compared on a ratio scale. Such comparisons of coefficients for single-item measures (reported elsewhere²⁸), and scale measures matched in terms of construct clearly indicated noteworthy proportional increases in true score variance for scale measures. Reliability coefficients for scales represented an increase of approximately 8 percent in true score variance over the reliability of items for the least reliable and shortest scale and an increase of approximately 57 percent for the most reliable and longest scale.

Test-retest reliability coefficients for scales ranged from a low of 0.41 to a high of 0.76 in East St. Louis and from a low of 0.59 to a high

of 0.86 in Sangamon County (see Table 5). Differences between these coefficients and internal-consistency reliability estimates were equal to or greater than 10 percent for five of the eight scales (the test-retest coefficients were lower in all instances where such differences were observed). For four of these five scales, differences between internal-consistency and test-retest coefficients of 10 percent or more were observed only in the East St. Louis field test.

Stability of Health Perceptions

Intertemporal stability coefficients (product-moment correlations between scores obtained two years apart) were positive and significant ($p < 0.001$, one-tailed test) for the four HPQ scales studied in the Tri-County field test. Coefficients, which may be interpreted as the proportion of variance that remained stable, were 0.31 for Health Worry/Concern, 0.45 for Resistance-Susceptibility to Illness, 0.59 for Prior Health, and 0.62 for Current Health. These coefficients may have been underestimated slightly due to noncomparability of alternate forms (Form I versus Form II of the HPQ) used to compute scores.²⁸

Validity: Higher Order Factor Structure

The three higher order factors explained approximately 70 percent of the total reliable variance in the eight HPQ scales in four field tests. In all field tests, a general factor accounting for slightly less than half of the reliable variance was observed in the unrotated solution. High loadings on this factor were observed for the Current Health, Prior Health, Resistance-Susceptibility to Illness, Health Outlook, Health Worry/Concern, and Sickness Orientation scales. These results suggest that scales in the first group share common variance, presumably due to favorable versus unfavorable perceptions of general health. As hypothesized, scales in the second group (Rejection of Sick Role and Attitude Toward Going to the Doctor) were unrelated to the unrotated general health factor.

The factors underlying these relationships were further clarified by the rotated higher order factor structure. Results are summarized across field tests in Table 6. It is clear from the rotated higher order factor structure that the eight HPQ scales fall into three groups in terms of

construct validity. The majority of reliable variance in scores for the two scales in the first group, Current Health and Prior Health, was accounted for by Present/Prior Health (Factor I). These scales differed in that only Current Health overlapped with Future Health (Factor II) and that approximately 33 percent of the reliable variance for Prior Health was not accounted for by the three higher order factors. Thus, high scores for Current Health reflect a favorable perception of present and prior health and to a significant but lesser extent, a favorable health outlook.

Approximately 30 to 50 percent of the reliable variance in scores for the four scales in the second group (Resistance-Susceptibility to Illness, Health Outlook, Health Worry/Concern, and Sickness Orientation) was accounted for by Factor II. It should be noted that while these four scales were similar in terms of their construct validity in relation to Factor II, important differences in terms of other sources of reliable variance were observed. Reliable variance in Resistance-Susceptibility to Illness and Health Outlook scale scores was also explained by Factor I, whereas variance in scores for the other two scales in this group (Health Worry/Concern and Sickness Orientation) was not. Further, these latter two scales differed with respect to overlap with Sick Role Propensity (Factor III). Variance in scores for the Sickness Orientation scale was not accounted for by Factor III, and approximately 65 percent of the reliable variance in that scale was not accounted for by the three higher order factors. For the Health Worry/Concern scale, approximately 31 percent of the reliable variance was accounted for by Factor II, approximately 14 percent by Factor III, and approximately 43 percent was not accounted for by the three higher order factors. Thus, Health Worry/Concern scores reflect health outlook but not current or prior health, and persons with high Health Worry/Concern scores were inclined toward accepting the sick role.

The remaining two scales, Rejection of Sick Role and Attitude Toward Going to the Doctor, were very similar in terms of known sources of variance. Both scales were essentially independent of Factors I and II, and a large amount of their reliable variances (54 and 64 percent, respectively) was explained by Factor III. A noteworthy amount of their reliable variances was not accounted for by the three higher order factors (see Table 6).

Validity: Health-Related Variables

Results regarding the validity of HPQ scales in relation to other measures of health status, health and illness behavior, and age are summarized in Tables 7-9. For this summary, three groups of scales have been formed on the basis of the higher order factor structure defined in Table 6.

Factor I Scales: Present/Prior Health. For both Current Health and Prior Health scales, relationships with other health variables were consistently significant and positive for other variables defining favorable health states (e.g., psychological well-being) and negative for other variables defining poor health (e.g., bed days) as hypothesized (see Table 7). Many of these coefficients were moderately high (i.e., above 0.40), indicating substantial relationships (not corrected for attenuation due to lack of perfect reliability). For both scales, correlations with health variables tended to be higher than correlations with health and illness behavior. For Current Health, correlations with other health variables tended to be higher for variables pertaining to feelings or perceptions (e.g., general health ratings, pain, and worry) than for more directly observable phenomena (e.g., bed days and role activity limitations).

Significant negative relationships between both Current Health and Prior Health and illness behaviors were observed as hypothesized; they tended to be higher for Current Health. Current Health, but not Prior Health, was also significantly related to compliance.

Current Health and Prior Health scale scores tended to decrease significantly with age in all field tests, as hypothesized.

Factor II Scales: Future Health. With very few exceptions, HPQ scales primarily measuring Factor II were consistently correlated with other health variables as hypothesized (see Table 8). Correlations tended to be more consistently significant and higher in absolute magnitude for Health Outlook and Resistance-Susceptibility to Illness than for Health Worry/Concern and Sickness Orientation. Significant coefficients involving the latter two scales and other health variables defined negligible relationships in many instances. All four of the Factor II scales appeared to be more related to whether respondents reported any bed days than to the number of bed days reported (for

those that reported one or more). It is interesting that the latter pattern was not observed for the two Factor I scales.

Resistance-Susceptibility to Illness, Health Outlook, and Health Worry/Concern were most often correlated with illness behavior as hypothesized; Sickness Orientation was not. Only Health Worry/Concern was significantly correlated with compliance (those who tended to worry more were more likely to comply).

Consistent with the validity hypothesis, Health Outlook scores decreased significantly with age in all field tests. Scores for the other three Factor II scales were related to age in the hypothesized direction when significant, although coefficients were not consistently significant across field tests.

Factor III Scales: Sick Role Propensity: The Rejection of Sick Role and Attitude Toward Going to the Doctor scales were not consistently correlated with the health variables studied (see Table 9); correlations that were significant (e.g., with pain and health worry) defined negligible relationships in most instances. The strongest relationships were between Attitude Toward Going to the Doctor and bed days (a behaviorally defined health variable) for those respondents reporting one or more bed days, and the general health item. The significant correlations between Factor III scales and behaviorally defined health variables were hypothesized; the correlation with the general health rating item was not. The absence of consistent and substantial relationships between Factor III scales and the health variables studied is consistent with the hypothesis that the Factor III scales do not measure health.

Although correlations between Factor III scales and health and illness behavior were often significant as hypothesized, they tended to define small or negligible relationships and they were not consistently significant. Those with a favorable attitude toward going to the doctor tended to report more recent doctor visits, check-ups, and compliance. Those tending to reject the sick role reported fewer hospitalizations (in two of three field tests).

As hypothesized,²⁷ tendencies to reject the sick role increased significantly with age in all field tests. Attitude toward going to the doctor was unrelated to age.

IV. DISCUSSION

The goals of this research were to develop and validate standardized survey measures of "perceived health".^{27,28,36} The conceptual orientation was toward general health ratings and personal assessments as opposed to specific components of health--physiologic, physical, mental, or social--and as opposed to directly observable phenomena such as days in bed due to poor health. In order to increase the amount of information obtained with general health rating instruments, items and scales were constructed to tap both positive and negative ends of the health continuum.

Early in the research,²⁷ it appeared that general health perceptions are oriented with respect to time, i.e., that respondents are able to distinguish among perceptions of what their general health has been, what it is now, and what they think it will be in the future. Further, it appeared that sick role propensity, a factor involved in conscious decisionmaking regarding sick role behavior, varies independently of general health perceptions. On the strength of these early study findings, a taxonomy of general health perceptions was defined, and a 36-item survey instrument (Form II of the Health Perceptions Questionnaire) was constructed and hypothesized to measure eight perceptual dimensions of general health and sick role propensity.

Traditional scaling methods were used to verify item groupings hypothesized to measure each of the eight dimensions. Thus, item groupings had to satisfy both theoretical and empirical criteria. Further, it should be noted that when the HPQ scales were constructed, scaling decisions were not based on any particular criterion of health. Because demonstrably valid measures of health perceptions did not exist, the taxonomy of health perceptions was formulated during a cyclical process of studying interrelationships among items measuring perceptions, formulation of hypotheses, and further studies of interrelationships. Emphasis was placed on evaluating scaling decisions and testing hypotheses in five populations in order to increase confidence in the generalizability of results. This procedure is in contrast to purely empirical approaches not based on specific hypotheses and without cross-validation such as in the construction of an index or scale on the basis of a single field test.

Scale Construction and Reliability of Scores

Attempts to construct scales from revised HPQ items were very successful. Only four of 36 items had to be discarded for failing to meet the stringent scaling criteria employed in four independent field tests. Retained items in hypothesized groupings were clearly shown to primarily measure the same constructs. On the basis of interpretation of the manifest content of the items in each group, the scales were given the following names: Current Health, Prior Health, Health Outlook, Sickness Orientation, Resistance-Susceptibility to Illness, Health Worry/Concern, Rejection of the Sick Role, and Attitude Toward Going to the Doctor. The fact that item groupings as hypothesized were empirically validated across independent field tests constitutes strong support for the taxonomy of health perceptions on which questionnaire construction was based. Furthermore, the consistency of findings across populations differing widely in age, level of educational attainment, income, and race constitutes strong support for the generalizability of conclusions regarding the item groupings and the scales they define.

Goals of the research were to achieve distributions of scale scores that have sufficient variance to warrant their use in general populations and that are roughly symmetrical, if not normally distributed, to facilitate their use in conjunction with the more powerful statistical methods. It was also believed that achievement of these and other goals, to be discussed below, were essential to the use of the HPQ as a *general* health measure. These goals were reached by questionnaire item revisions designed to achieve roughly symmetrical distributions of responses to items in conjunction with careful construction of scales from items. It is suggested that these methods be applied to the measurement of other health phenomena for which a continuum of ordinal or interval scores is appropriate.

It was clearly shown that scale scores are sufficiently reliable for purposes of group comparisons and that single-item scores often are not. Given that the great majority of studies of general health ratings published in the last 25 years have been based on single-item measures,¹ and given the poor reliability of single-item measures (particularly in studies involving disadvantaged respondents), it is likely that the strength of most associations between general health perceptions and other variables, as reported in the published literature, has been underestimated.

Internal-consistency and test-retest reliability coefficients for HPQ scales tended to be lower in disadvantaged populations, although the amount of true score variance defined by scales was sufficient to warrant their use in all populations studied. These results are impressive in view of the fact that the eight HPQ scales are very short, containing only two to nine items each, and that the 32 items necessary to construct the eight scales require only about seven minutes to administer, on the average. When reliability coefficients for disadvantaged and non-disadvantaged respondents were compared, differences between test-retest coefficients were considerably greater than differences between internal-consistency coefficients. These trends suggest that general health perceptions are less reliable at a point in time *and* less stable over time for the disadvantaged.

The four HPQ constructs (Current Health, Prior Health, Resistance-Susceptibility to Illness, and Health Worry/Concern) that were evaluated in terms of stability over a two-year period in one field test site were shown to be sufficiently stable to warrant their use in repeated-measures research designs. Specifically, the use of general health ratings in a repeated-measures design would improve precision of estimates (or would permit a reduction in the sample size necessary to detect group differences) for purposes of hypothesis testing. The finding that general health perceptions tend to be stable over time is consistent with published results;^{9,12,17} however, current study findings suggest that the long-term stability of these perceptions has been underestimated. Most likely, the lower estimates in published studies resulted from the lower reliability (at each point in time) of the single-item scores that were used. Favorable results regarding the reliability and stability of general health ratings are particularly noteworthy given the prejudice that exists against use of such ratings on these grounds.

Validity of Scale Scores

Demonstration that the items used to construct a scale (or index) satisfy the assumptions of the scaling methods used and that resulting scores are reliable is an essential prerequisite to use of the scale for any purpose. However, it is not sufficient to know that a score is reliable, or to infer score meaning entirely from face validity, i.e., from

the manifest content of questionnaire items. It is essential that empirical methods be used to determine the validity of scale scores so that their appropriate uses as measures of health can be established. It is also important to study interrelationships among scales in a battery in order to determine if they are all necessary, i.e., if each scale constructed from a questionnaire contributes information in addition to that provided by the others. These important issues were addressed by studying interrelationships among HPQ scales and relationships between scales and a range of health-related variables.

Relationships Among HPQ Scales. Results of empirical studies of relationships among the eight HPQ scales clearly indicate that each scale score primarily reflects one of the three kinds of perceptions that were identified: past/present health, future health, or sick role propensity. When the eight scales were grouped in terms of which of the three dimensions accounted for the majority of reliable variance in their scores, the results were amazingly consistent across field tests. These results are important for several reasons.

First, the results constitute strong support for the construct validity of HPQ scale scores, i.e., each scale tends to measure the perceptual construct it was intended to measure to a greater extent than it measures constructs that it was not intended to measure. The findings constitute strong support for construct validity because the results were very clear for most scales and because the results were nearly identical across independent field tests.

Second, the results are an important first step in the difficult process of determining the meaning that should be assigned to scores computed from these measures and, therefore, the appropriate uses of the scales. In this regard, it is recommended that Table 6 be used as an adjunct to other validity findings in interpreting HPQ scale scores. Factor analytic evidence of construct validity also served to call attention to the fact that the HPQ scales, as do most health measures, tap more than one health construct. Although each HPQ scale consistently had a major source of reliable variance (either perceptions of present/prior health, future health, or sick role propensity), none of the scales had a single source of reliable variance. For example, Current Health scores indicate favorable perceptions of health in the present

and, to a lesser but noteworthy extent, a favorable prior health history and health outlook. Health Worry/Concern scores reflect both an unfavorable perception of health in the future and a greater propensity to accept the sick role.

Third, factor analytic studies of construct validity are useful in determining the extent to which gains in efficiency with use of a smaller number of scale scores (based on more global concepts) are warranted given the loss of some reliable information. In the case of the HPQ, findings suggest that three global scale scores corresponding to perceptions of past/present health, future health, and sick role propensity would capture the majority of reliable variance defined by the battery. However, when the same findings are viewed from another perspective, it is clear that a noteworthy amount of reliable variance (one-third to two-thirds in the case of four scales) would be lost. Further research is necessary to better understand the practical implications of these tradeoffs.

Other Health Variables. It was hypothesized that six of the HPQ scales (Current Health, Prior Health, Health Outlook, Resistance-Susceptibility to Illness, Sickness Orientation, and Health Worry/Concern) would measure health and that two scales would not (Rejection of Sick Role and Attitude Toward Going to the Doctor). Findings regarding the relationships among the eight scales and higher order factors pertaining to health and sick role propensity constitute support for these hypotheses. However, such evidence is not sufficient. If the scales in the first group are valid measures of general health they should be substantially related to other health variables including those that tap physical and mental health components. Further, scales in the second group should not be related to health variables (except possibly the behaviorally defined variables, as hypothesized) and should be related to health and illness behavior.

Relationships between the six HPQ scales hypothesized to measure health and the 11 other health variables studied were strong enough to conclude that the scales measure health and weak enough to indicate that the scales contribute unique information about health. Health perceptions tended to be more unfavorable in conjunction with increases in: role activity limitations, bed days, sickness, chronic health problems, pain, and worry. Health perceptions also tended to go hand in hand with psychological well-

being and, to a marked degree, with favorable health ratings obtained using a different method. Current Health scores were substantially related to both physical health variables (e.g., chronic problems and role activity limitations) and mental health variables (e.g., psychological well-being and worry). These results constitute strong support for the validity of HPQ scales as general measures of health.

It is important to keep in mind that, while the current research has contributed to increased understanding of the meaning of general health ratings, little or nothing is known about the clinical significance of these scores. Data from the current study, particularly when nonlinear relationships with variables defining poor health were examined,²⁸ and data from other sources^{37,38} suggest that the majority of very low general health rating scores (i.e., those in the lowest decile of the score distribution) can be explained in terms of chronic functional limitations due to poor health, chronic diseases, and psychiatric impairment. Furthermore, correlations with health measures that tap positive states (e.g., vigor and happiness) suggest that very high scores can be explained by states of positive well-being.³⁸ However, considerable effort beyond the current research would be necessary to determine what, if any, clinical significance can be assigned to general health rating scores.

Health and Illness Behavior. Studies of the validity of HPQ scales in relation to health and illness behavior employed cross-sectional research designs. Ratings of health and sick role propensity (at a given point in time) were correlated with information obtained at the same time regarding *prior* health and illness behavior. These data were used to test validity hypotheses regarding the effects of health perceptions and sick role propensity on subsequent behavior. Given that the prior behaviors are certain to have affected present ratings,^{28,39} and the effects of other methodological problems,⁴⁰ these methods permit only weak tests of the validity of the ratings in relation to behavior. Despite these limitations, validity hypotheses regarding HPQ health scales and health and illness behavior were confirmed, supporting the validity of the scales as measures of health. Correlations between illness behaviors and most HPQ health scales were consistently significant in the hypothesized direction and the correlations were often substantial (e.g., Current Health in relation to doctor visits). Findings regarding the negative relationship

between general health perceptions and use of health care services are consistent with those reported by others.^{6,16,20} Only Current Health and Health Worry/Concern were significantly related to compliance with medical regimens. Consistent with previous findings^{20,27} and current study hypotheses, health perceptions were not related to health behavior (e.g., check-ups). It should be noted that Current Health and Prior Health scales were more strongly and more consistently correlated with the illness behaviors studied than were other HPQ health scales (e.g., Health Outlook). This inconsistency may be an artifact of the cross-sectional study design. A test of the validity of the HPQ scales with a prospective study design is underway pursuant to the Health Insurance Study which The Rand Corporation is conducting for the U.S. Department of Health, Education, and Welfare.⁴¹

Hypothesized relationships between HPQ scales pertaining to sick role propensity and health and illness behavior were confirmed in many instances. However, correlations often defined negligible relationships and were inconsistent across field tests in some instances. Attitude Toward Going to the Doctor tended to be more valid than Rejection of the Sick Role in relation to both the health and illness behaviors studied. These results offer weak support for the validity of the two scales and suggest that a more accurate interpretation of their scores is that they tap "patient role" propensity (as opposed to "sick role" propensity).

It is important to note that correlations among the HPQ scales were not taken into account in evaluating their validities in relation to behavior. Thus, it is not possible to determine from these results whether or not all of the scales make a unique contribution to an explanatory model. This important issue is addressed elsewhere.²⁸ Briefly, several of the scales make a unique contribution to the explanation of health and illness behavior. When health and illness behaviors were regressed on the eight HPQ scales in order to take correlations among the latter into account in two field tests, the scales explained approximately 15 to 20 percent of the variance in doctor visits, 2 to 6 percent of the variance in check-ups, and 3 percent of the variance in dental visits. Only three or four scales, but not the same scales, were necessary to explain all of the variance that could be explained in the three behaviors studied.

Health Perceptions and Age. In support of their validities as measures of health, Current Health, Prior Health, and Health Outlook scale scores were significantly negatively correlated with age in all field tests. These results are consistent with those previously reported.^{8,10,21,22,27} When examined for nonlinearity,²⁸ relationships between age and general health perceptions tended to be linear, when significant. When curvilinear trends were observed, they were weak and due to the fact that very favorable health perceptions were only rarely observed for older age groups (above 70), whereas the full range of favorable and unfavorable scores was observed for younger respondents. Age relationships with other HPQ health scales were not consistent. Significant positive correlations between age and Rejection of Sick Role observed in all field tests suggest that, in addition to perceiving their health as poorer, older (more than younger) persons tend to consciously resist letting illness interfere with their lives. The latter findings are consistent with those reported for Form I of the HPQ.²⁷

V. CONCLUSIONS

Considerable progress has been made regarding the conceptualization and measurement of general health perceptions. Eight scales that are reliable and stable in diverse populations have been constructed from 32 items in a standardized survey instrument that can be self-administered in about seven minutes. While further research is necessary to understand the clinical significance of scale scores, available information regarding validity will be useful in interpreting scores and in establishing the appropriate uses of the scales. In that regard, it appears that the scales should be used whenever general health measures, i.e., those tapping both subjective and objective health constructs along the continuum from negative to positive well-being, are desired. For example, these measures would contribute unique information in evaluations of medical care services, in studies designed to explain health and illness behavior, in studies of relationships among health constructs, and in population assessments of general health status. Further research should focus on the development and validation of shorter forms of the scales, instruments constructed specifically for use with children, and on the validity of the scales in prospective study designs. The latter is necessary to clarify the causes and effects that are operating in relationships among these and other health constructs as well as health and illness behavior.

Table 1
HEALTH PERCEPTION QUESTIONNAIRE ITEMS AND HYPOTHESIZED ITEM
GROUPINGS, FORM II^a

Hypothesized _b Item Grouping	Item
CH	According to the doctors I've seen, my health is now excellent.
SR	I try to avoid letting illness interfere with my life.
RS	I seem to get sick a little easier than other people.
HO	I think my health will be better in the future than it is now. ^c
SR	When I'm sick I try to keep it to myself.
PH	I have been in bed a lot in the past because of illness. ^c
CH	The people I know seem to be healthier than I am. ^c
SR	When I'm sick I try to just keep going as usual.
HO	I will probably be sick a lot in the future.
WC	I never worry about my health.
CH	I feel better now than I ever have before.
SO	Getting sick once in a while is a part of my life.
WC	I worry about my health more than other people worry about their health.
RS	Most people get sick a little easier than I do.
CH	I am somewhat ill.
RS	My body seems to resist illness very well.
HO	In the future, I expect to have better health than other people I know.
CH	I'm not as healthy now as I used to be.
PH	I was so sick once I thought I might die.
AD	I don't like to go to the doctor.
WC	My health is a concern in my life.
CH	I'm as healthy as anybody I know.
HO	I expect to have a very healthy life.
AD	It doesn't bother me to go to a doctor.
CH	My health is excellent.
PH	I've never had an illness that lasted a long period of time.
HO	Most of the people I know will probably have fewer health problems than I will in the future. ^c

Hypothesized ^b Item Grouping	Item
WC	Others seem more concerned about their health than I am about mine.
SO	I accept that sometimes I'm just going to be sick.
CH	I have been feeling bad lately.
PH	I have never been seriously ill.
RS	When there is something going around I usually catch it.
SR	When I think I am getting sick, I fight it.
CH	Doctors say that I am now in poor health.
HO	I think my health will be worse in the future than it is now.
CH	I feel about as good now as I ever have.

^aItems are listed in order of administration.

^bCH = Current Health; PH = Prior Health; HO = Health Outlook;
WC = Health Worry/Concern; RS = Resistance-Susceptibility;
SO = Sickness Orientation; SR = Rejection of Sick Role;
AD = Attitude Toward Going to the Doctor.

^cThese items did not satisfy scaling criteria and therefore were not retained.

Table 2
SUMMARY OF RESPONDENT CHARACTERISTICS, FIVE FIELD TESTS

Characteristics	Field Tests ^a				
	ESL	SC	FP	LAC	TC
Sample Size	323	432	527	640	92
Sex					
Male (%)	19	22	36	63	16
Female (%)	81	78	64	37	84
Race					
White (%)	10	97	b	65	90
Nonwhite(%)	90	3	b	35	10
Age					
Minimum	17	17	17	18	20
Maximum	88	84	84	92	80
Median	43	45	32	43	51
Family Income (\$)					
Minimum	0	<2,000	0	0	0
Maximum	20,000+	20,000+	20,000+	30,000+	20,000
Median	5,400	11,900	12,000	9,500	9,000
Education (yrs.)					
Minimum	3	3	6	0	5
Maximum	20+	20+	20+	20+	17
Median	11	12	14	12	11

^aESL = East St. Louis
SC = Sangamon County
FP = Family Practice
LAC = Los Angeles County
TC = Tri-County

^bNot available

Table 3

DEFINITION OF VALIDITY VARIABLES AND DIRECTION OF
HYPOTHEZED RELATION WITH HEALTH

Category/Variable	Definition	Hypothesized Relationship with Health
<u>Health Status</u>		
General health item	Rating of general health in terms of excellent, good, fair, or poor	+
Role activity limitations	Number of days during the prior year that the respondent was too sick to perform usual activities	-
Any bed days	Whether respondent spent all or part of any day in bed because of illness or injury during the prior two months	-
Total bed days	For those respondents with one or more bed days during the prior two months, the number of days in which all or part of the day was spent in bed because of illness or injury	-
Sickness	Presence or absence of sickness during the prior two months	-
Chronic problems	Presence or absence of any continuous or recurring health problem (one that is present all the time or that makes the respondent ill from time to time)	-
Pain	Four-choice rating of how often pain was experienced during the prior year	-
Feeling poorly	Number of days during the prior two months on which the respondent did not feel as well as usual	-
Health worry	Four-choice rating of worry about health during the prior year	-

Table 3 (continued)

Category/Variable	Definition	Hypothesized Relationship with Health
<u>Health Status</u> (continued)		
Psychological well-being	10-item scale to measure positive versus negative psychological well-being [42]	+
Reason for last doctor visit	Whether the last doctor visit was for sickness/injury	-
<u>Health and Illness Behavior</u>		
Number of doctor visits	Number of doctor visits during the prior year	-
Dental visit	Whether the respondent saw a dentist during the prior year	-
Check-up	Whether a doctor was seen for a check-up or physical examination, when not sick, during the prior year	0
Doctor visit when needed	For those respondents reporting sickness, injury, or accident during the prior two months, whether a doctor was seen	-
Recency of care	Number of months since last doctor visit for any reason	+
Hospitalization	Whether respondent stayed overnight in a hospital (other than for childbirth) during the prior year	-
Compliance	Whether the respondent had been following doctor's orders exactly, for those who were given medical regimens	-

Table 4
MEANS AND STANDARD DEVIATIONS (IN PARENTHESES) FOR HPQ FORM II
SCALES, FOUR FIELD TESTS

Scales	k ^a	Highest Possible Score	Field Tests ^b			
			ESL	SC	FP	LAC
Current Health	9	45	27.6 (8.5)	32.9 (7.3)	32.7 (7.9)	32.7 (7.8)
Prior Health	3	15	9.4 (3.6)	10.5 (3.4)	10.5 (3.8)	10.3 (3.6)
Resistance-Susceptibility	4	20	13.8 (3.0)	15.0 (2.6)	14.3 (3.2)	15.3 (2.6)
Health Outlook	4	20	13.1 (2.4)	14.1 (2.6)	14.2 (2.8)	14.3 (2.8)
Health Worry/Concern	4	20	13.4 (2.8)	11.4 (2.7)	12.1 (2.8)	11.6 (3.1)
Sickness Orientation	2	10	7.1 (1.9)	6.7 (1.9)	7.0 (1.8)	6.2 (2.2)
Rejection of Sick Role	4	20	11.4 (3.0)	11.5 (2.3)	10.9 (2.5)	11.3 (2.5)
Attitude Toward Going to the Doctor	2	10	5.8 (2.3)	5.8 (2.1)	6.1 (2.4)	6.3 (2.2)

^aNumber of items in scale; also equals lowest possible score

^bESL = East St. Louis
SC = Sangamon County
FP = Family Practice
LAC = Los Angeles County

Table 5
Internal-Consistency and Test-Retest Reliability Coefficients for
Eight HPQ Scales

	k ^a	Internal-Consistency Coefficients ^b			Test-Retest Coefficients ^c	
		Low	High	Median	ESL	SC
Current Health	9	.89	.92	.91	.76	.86
Prior Health	3	.70	.79	.73	.67	.78
Health Outlook	4	.64	.79	.75	.54	.76
Resistance- Susceptibility	4	.58	.80	.71	.74	.73
Health Worry/ Concern	4	.45	.62	.60	.60	.65
Sickness Orien- tation	2	.46	.61	.59	.42	.72
Rejection of Sick Role	4	.54	.60	.59	.41	.66
Attitude Toward Going to Doctor	2	.62	.79	.67	.51	.59

^ak = Number of items used to compute scale score.

^bSummary of coefficients across four field tests.

^cESL = East St. Louis
SC = Sangamon County

Table 6

PERCENTAGES OF RELIABLE VARIANCE IN EIGHT HPQ SCALES ACCOUNTED AND NOT ACCOUNTED FOR BY HIGHER ORDER FACTORS^a

Scales	Reliable Variance	Sources of Reliable Variance ^b			Unknown
		Present/Prior Health (Factor I)	Future Health (Factor II)	Sick Role Propensity (Factor III)	
Current Health	91	68	19	01	10
Prior Health	74	62	03	00	33
Resistance-Susceptibility	70	18	54	06	19
Health Outlook	73	26	45	03	20
Health Worry/Concern	55	05	31	14	43
Sickness Orientation	55	02	29	00	65
Rejection of Sick Role	58	02	02	64	29
Attitude Toward Going to Doctor	69	00	03	54	38

^aTable entries are averages across four field tests; decimal points have been omitted.

^bSources of reliable variance do not total 100 by as much as 2 to 7 percent due to errors of estimation.

Table 7

Summary of Validity Coefficients for Current Health and Prior Health
Scales in Relation to Health Status, Health
and Illness Behavior, and Age

Validity Variables	Factor I Scales			
	Current Health		Prior Health	
	# ^a	r ^b	# ^a	r ^b
<u>Health Status</u>				
General Health Item	1/1	.80 [*]	1/1	.50 [*]
Role Activity Limitations	4/4	-.46 [*]	4/4	-.36 [*]
Any Bed Days	1/1	-.37 [*]	1/1	-.21 [*]
Total Bed Days	1/1	-.38 [*]	1/1	-.30 [*]
Sickness	1/1	-.24 [*]	1/1	-.18 [*]
Chronic Problems	1/1	-.43 [*]	1/1	-.32 [*]
Pain	1/1	-.55 [*]	1/1	-.39 [*]
Feeling Poorly	1/1	-.28 [*]	1/1	-.19 [*]
Health Worry	1/1	-.58 [*]	1/1	-.44 [*]
Psychological Well-Being	1/1	.39 [*]	1/1	.19 [*]
Reason for Doctor Visit	1/1	.21 [*]	1/1	.11 [*]
<u>Health and Illness Behavior</u>				
Number of Doctor Visits	3/3	-.42 [*]	3/3	-.32 [*]
Dental Visits	1/3	.03	2/3	-.02
Check-up	1/3	-.05	0/3	-.03
Doctor Visit When Needed	1/1	-.47 [*]	1/1	-.27 [*]
Recency of Care	2/2	.16 [*]	2/2	.15 [*]
Hospitalization	3/3	-.30 [*]	3/3	-.28 [*]
Compliance	1/1	-.14 [*]	0/1	-.07
<u>Demographic Variable</u>				
Age	4/4	-.26 [*]	4/4	-.19 [*]

^aNumber of significant coefficients in the hypothesized direction, in relation to number of field tests.

^bProduct-moment correlation not corrected for attenuation. In those instances where more than one correlation was independently computed, the table entry is the median coefficient across fieldtests.

Table 8

Summary of Validity Coefficients for Resistance-Susceptibility, Health Outlook, Health Worry/Concern and Sickness Orientation Scales in Relation to Health Status, Health and Illness Behavior, and Age

Validity Variables	Factor II Scales							
	Resistance Susceptibility		Health Outlook		Health Worry/Concern		Sickness Orientation	
	# ^a	r ^b	# ^a	r ^b	# ^a	r ^b	# ^a	r ^b
<u>Health Status</u>								
General Health Item	1/1	.50*	1/1	.58*	1/1	-.38*	1/1	-.16*
Role Activity Limitations	4/4	-.33*	4/4	-.24*	4/4	.19*	4/4	.14*
Any Bed Days	1/1	-.24*	1/1	-.23*	1/1	.10*	1/1	.14*
Total Bed Days	0/1	.00	0/1	-.16	0/1	.07	0/1	.01
Sickness	1/1	-.21*	0/1	-.13	1/1	.12*	1/1	.12*
Chronic Problems	1/1	-.26*	1/1	-.34*	1/1	.10*	1/1	.14*
Pain	1/1	-.41*	1/1	-.35*	1/1	.24*	1/1	.18*
Feeling Poorly	1/1	-.21*	1/1	-.13*	1/1	.14*	1/1	.09*
Health Worry	1/1	-.42*	1/1	-.33*	1/1	.42*	1/1	.20*
Psychological Well-Being	1/1	.18*	1/1	.28*	1/1	-.12*	0/1	-.11
Reason for Doctor Visit	1/1	.17*	1/1	.09*	0/1	-.05	0/1	.01
<u>Health and Illness Behavior</u>								
Number of Doctor Visits	3/3	-.32*	3/3	-.16*	3/3	.20*	1/3	.10*
Dental Visits	1/3	-.02	1/3	.06	0/3	-.01	0/3	.00
Check-up	0/3	-.01	1/3	-.01	0/3	-.04	0/3	.02
Doctor Visit When Needed	1/1	-.21*	1/1	-.26*	1/1	.17*	0/1	.09
Recency of Care	2/2	.15*	1/2	.06	2/2	-.15*	0/2	.00
Hospitalization	3/3	-.26*	2/3	-.11*	1/3	-.07	0/3	.02
Compliance	0/1	-.06	0/1	-.06	1/1	.16*	0/1	-.10
<u>Demographic Variable</u>								
Age	1/4	.09	4/4	-.23*	3/4	-.10*	2/4	-.06

^aNumber of significant coefficients in the hypothesized direction, in relation to number of field tests.

^bProduct-moment correlation not corrected for attenuation. In those instances where more than one correlation was independently computed, the table entry is the median coefficient across fieldtests.

Table 9

Summary of Validity Coefficients for Rejection of Sick Role and Attitude Toward Going to Doctor Scales in Relation to Health Status, Health and Illness Behavior, and Age

Validity Variables	Factor III Scales			
	Rejection of Sick Role		Attitude Toward Going to Doctor	
	#a	r ^b	#a	r ^b
<u>Health Status</u>				
General Health Item	0/1	.02	1/1	.23*
Role Activity Limitations	0/4	-.03	0/4	-.03
Any Bed Days	0/1	.01	0/1	-.03
Total Bed Days	0/1	.06	1/1	.22*
Sickness	0/1	-.04	0/1	-.03
Chronic Problems	1/1	-.08*	0/1	.00
Pain	1/1	-.11*	1/1	-.13*
Feeling Poorly	0/1	-.03	0/1	-.07
Health Worry	1/1	-.12*	1/1	-.13*
Psychological Well-Being	0/1	.04	0/1	.05
Reason for Last Doctor Visit	0/1	.00	0/1	.07
<u>Health and Illness Behavior</u>				
Number of Doctor Visits	0/3	-.02	1/3	.08
Dental Visits	0/3	-.01	1/3	.09
Check-up	0/3	.00	2/3	.15*
Doctor Visit When Needed	0/1	.01	0/1	.00
Recency of Care	0/2	-.03	2/2	-.11*
Hospitalization	2/3	-.10*	0/3	.00
Compliance	0/1	-.03	1/1	.21*
<u>Demographic Variable</u>				
Age	4/4	.14*	0/4	.04

^aNumber of significant coefficients in the hypothesized direction, in relation to number of field tests.

^bProduct-moment correlation not corrected for attenuation. In those instances where more than one correlation was independently computed, the table entry is the median coefficient across fieldtests.

REFERENCES

1. Donald, C.A. and J.E. Ware, Jr. A review of the literature on general health ratings. Santa Monica, California: The Rand Corporation, P-5721, forthcoming.
2. Berkman, P. Life stress and psychological well-being: A replication of Langner's analysis in the Midtown Manhattan Study. *J Health Soc Behav* 12:35, March 1971.
3. Hochstim, J.R., D.A. Athanasopoulos and J.H. Larkins. Poverty area under the microscope. *Am J Public Health* 58:1815, October 1968.
4. Madow, W.G. Interview data on chronic conditions: Compared with information from medical records. *Vital and Health Statistics*, Series 2, No. 23, May 1967.
5. Edwards, J. and D. Klemmack. Correlates of life satisfaction: A reexamination. *J Gerontol* 28:497, 1973.
6. Fabrega, H. and R.E. Roberts. Social-psychological correlates of physician use by economically disadvantaged negro urban residents. *Med Care* 10:215, May-June 1972.
7. Reynolds, J., W.A. Rushing and D.L. Miles. The validation of a Function Status Index. *J Health Soc Behav* 15:271, 1974.
8. Schnore, L.F. and J.D. Cowhig. Some correlates of reported health in metropolitan centers. *Soc Probs* 7:218, 1959-60.
9. Suchman, E.A., B.S. Phillips and G.F. Streib. An analysis of the validity of health questionnaires. *Soc Forces* 26:223, 1958.
10. Cole, S. and R. Lejeune. Illness and the legitimation of failure. *Am Soc Rev* 37:347, June 1972.
11. Cutler, S. Voluntary association participation and life satisfaction: A cautionary research note. *J Gerontol* 28:96, 1973.
12. Maddox, G.L. Self-assessment of health status: A longitudinal study of selected elderly subjects. *J Chron Disab* 17:449, 1964.
13. Renne, K.S. Correlates of dissatisfaction in marriage. *J Marriage Family* 22:54, February 1970.
14. Renne, K.S. Health and marital experience in an urban population. *J Marriage Family* 23:338, May 1971.
15. Renne, K.S. Measurement of social health in a general population survey. *Soc Sci Res* 3:25, 1974.
16. Friedsam, H.J. and H.W. Martin. A comparison of self and physicians' health ratings in an older population. *J Health Human Behav* 4:179, 1963.

17. Heyman, D.K. and F.C. Jeffers. Effect of time lapse on consistency of self-health and medical evaluations of elderly persons. *J Gerontol* 18:160, 1963.
18. Kessel, N. and M. Shepard. The health and attitudes of people who seldom consult a doctor. *Med Care* 3:6, 1965.
19. Ludwig, E.G. and G. Gibson. Self-perception of sickness and the seeking of medical care. *J Health Soc Behav* 10:125, June 1969.
20. Tissue, T. Another look at self-rated health among the elderly. *J Gerontol* 27:91, 1972.
21. Gaitz, C.M. and J. Scott. Age and the measurement of mental health. *J Health Soc Behav* 12:55, March 1972.
22. Palmore, E. and C. Luikart. Health and social factors related to life satisfaction. *J Health Soc Behav* 13:68, March 1972.
23. Spreitzer, E. and E.E. Snyder. Correlates of life satisfaction among the aged. *J Gerontol* 29:454, 1974.
24. Pratt, L. The relationship of socioeconomic status to health. *Am J Public Health* 61:281, February 1971.
25. MacPherson, A.S. *The Measurement of Mental Health in a General Population*. A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Science in Medicine. Hamilton, Ontario: McMaster University, 1972.
26. Stevens, S.S. Mathematics, measurement, and psychophysics. In S.S. Stevens (ed.), *Handbook of Experimental Psychology*, pp. 1-49. New York: John Wiley and Sons, 1951.
27. Ware, J.E., Jr., W.R. Wright and M.K. Snyder. *Measures of Perceptions Regarding Health Status: Preliminary Findings as to Scale Reliability, Validity and Administration Procedures*. Publication No. PB-242-726. Springfield, Va.: National Technical Information Service, 1974.
28. Ware, J.E., Jr. *Development and Validation of Scales to Measure Perceived Health*. Volume II of the Final Report on Contract No. HSM 110-72-299, prepared for the Research Methods Branch, National Center for Health Services Research, U.S. Department of Health, Education, and Welfare. Carbondale, Ill.: School of Medicine, Southern Illinois University, 1976.
29. Ware, J.E., Jr. and W.R. Wright. *Development and Validation of Scales to Measure Key Health Concepts: Methodological Appendix*. Volume IV of the Final Report on Contract No. HSM 110-72-299, prepared for the Research Methods Branch, National Center for Health Services Research, U.S. Department of Health, Education, and Welfare. Carbondale, Ill.: Southern Illinois University, 1976.
30. Comrey, A.L. *A First Course in Factor Analysis*. New York: Academic Press, 1973.

31. Comrey, A.L. Tandem criteria for analytic rotation in factor analysis. *Psychometrika* 32:143, June 1967.
32. Campbell, D.T. and F.W. Fiske. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psych Bull* 56:81, 1959.
33. Howard, K.I. and G.G. Forehand. A method for correcting item-total correlations for the effect of relevant item inclusion. *Educ Psych Meas* 22:731, 1962.
34. Cronbach, L.J. Coefficient Alpha and the internal structure of tests. *Psychometrika* 16:297, September 1967.
35. Helmstadter, G.C. *Principles of Psychological Measurement*, pp. 355-61. New York: Appleton-Century-Crofts, 1973.
36. Ware, J.E., Jr., M.K. Snyder, E.R. McClure and I.W. Jarret. *The Measurement of Health Concepts*. Publication No. PB-239-508/AS. Springfield, Va.: National Technical Information Service, December 1972.
37. Stewart, A., J.E. Ware, Jr. and R.H. Brook. The meaning of health: Understanding functional limitations. Santa Monica, Ca.: The Rand Corporation, P-5660, forthcoming.
38. Brook, R.H., J.E. Ware, Jr. et al. *Conceptualization and Measurement of Health for Adults in the Health Insurance Study*: Dayton. Santa Monica, Ca.: The Rand Corporation, R-1987-HEW, forthcoming.
39. Ware, J.E., Jr., M.K. Snyder and W.R. Wright. *Development and Validation of Scales to Measure Patient Satisfaction with Health Care Services*. Volume I of the Final Report on Contract No. HSM 110-72-299, prepared for the Research Methods Branch, National Center for Health Services Research, U.S. Department of Health, Education, and Welfare. Carbondale, Ill.: Southern Illinois University, 1976.
40. Ware, J.E., Jr. A cross-sectional study of patient satisfaction and use of health care services in rural southern Illinois. Santa Monica, Ca.: The Rand Corporation, P-5691, forthcoming.
41. Newhouse, J.P. A design for a Health Insurance Experiment. *Inquiry* 11:5, 1974.
42. Bradburn, N.M. *The Structure of Psychological Well-Being*. Chicago: Aldine Publishing Company, 1969.

