Scales for Measuring General Health Perceptions

by John E. Ware Jr.

This article reports on the construction and testing of eight health perception scales from 32 items on a standardized survey instrument designed for self-administration, the Health Perceptions Questionnaire (Form II). The scales measure perceptions of prior health, current health, health outlook, resistance/susceptibility to illness, health worry/concern, sickness orientation, rejection of sick role, and attitude toward going to the doctor. Field testing revealed that the scales are valid, reliable, and stable over time for diverse populations. It is recommended that the scales be used in studies requiring general health measures. Suggestions for future research are offered.

At the instigation of the National Center for Health Services Research (NCHSR), we have been working on the development of standardized measures that can be used to survey people's perceptions concerning their own health [1–3]. Our conceptual orientation in this work has been 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.

We developed a general health rating instrument, the Health Perceptions Questionnaire [1], and, on the basis of findings gained with this instrument, developed a taxonomy of general health perceptions and constructed a revised survey instrument (Health Perceptions Questionnaire, Form II). This instrument, which consists of 36 items structured as statements of opinion regarding health, measures eight perceptual dimensions of general health and sick role propensity: prior health, current health, health outlook, resistance/susceptibility to illness, health worry/concern, sickness orientation, rejection of sick role, and attitude toward going to the doctor [1].

This article describes the construction and testing of the eight scales of health perceptions and suggests uses for the scales. The specific goals of the research were to confirm the item groupings in Form II of the Health Perceptions Questionnaire (HPQ) hypothesized from the field test of Form I, to

Research supported by contract no. HSM 110-72-299, National Center for Health Services Research, DHEW, and the Health Insurance Study grant from DHEW.

Address communications and requests for reprints to John E. Ware Jr., Ph.D., Research Psychologist, Rand Corporation, 1700 Main Steet, Santa Monica, CA 90406.

evaluate the items in terms of scaling criteria to achieve roughly normally distributed continuous scale scores, to assess the reliability of item and scale scores and the long-term stability of scores, and to evaluate the validity of scale scores.

Although many published studies have included ratings of general health (for example, see refs. 4–28), psychometric evaluations of such ratings have not been performed. Published reports indicate confusion regarding the level of enumeration [29] achieved with general health ratings, and from the statistical methods used in the construction of scales it appears that the same or very similar general health rating items have been enumerated as nominal [7,11,12,14–17,19–22], ordinal [23], and interval [6,8–10,13,19,24–26] data. With few exceptions [5,12,18,27,28], ratings have been scored from responses to single questionnaire items in absence of evidence that single item ratings are reliable. We hope that the procedures used in the evaluation of the HPQ scales will be useful in other studies involving ratings of general health.

Methods

Data

Data were gathered both with and without interviewer supervision using Form II of the HPQ (see Table 1). The HPQ requires approximately eight minutes to administer, on the average. The entire interview schedule [30], including standardized survey instruments used to gather data regarding demographic and socioeconomic characteristics of the respondents and health-related information used in validity analyses, requires less than an hour to administer, on the average.

Analyses were based on data from five field tests involving approximately 2,000 adult respondents sampled from general populations between 1973 and 1975. The field tests were done in East St. Louis (field test A), Sangamon County, Illinois (field test B), a family practice in Springfield, Illinois (field test C), Los Angeles County (field test D), and Franklin, Perry, and Williamson Counties in southern Illinois (field test E). In all field tests, mixed sampling designs were used to draw representative households, and one or more adults from each household were selected [2].

In two of the field tests (C and E) questionnaires were mailed, and in these tests the return rate of complete and usable interviews was approximately 37 percent. In three field tests (A, B, and D) interviewers took questionnaires around to designated households and assisted respondents in filling them out; rates of usable returns were 62, 82, and 95 percent, respectively. In field test D, interviewers also simply left questionnaires with some randomly selected respondents and asked them to fill them out and mail them back; return rate of completed interviews was 67 percent using this method. The effects of data-gathering methods and partial returns on conclusions about health from these field tests are discussed in ref. 2. It was assumed that data-gathering methods

and differences in return rates did not constitute potential sources of bias in the conclusions drawn in this article.

Plan of Analysis

Scaling of Items. The HPQ items were first grouped according to the eight specific perceptions they were hypothesized to measure (see Table 1). The hypothesized item groupings were based on the results of factor analytic studies of Form I of the HPQ [1].

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 [31]. 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 [31,32].

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 validity [33] 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 Forehand [34]) 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 is accompanied by five response categories: definitely true, mostly true, don't know, mostly false, and definitely false. 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 from 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's [35] alpha coefficient.

The 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, we used two methods of validation. The first 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.

Table 1. Health Perception Questionnaire Items and Hypothesized Item Groupings, Form II

Item grouping	Item
Current health	According to the doctors I've seen, my health is now excellent
	The people I know seem to be healthier than I am*
	I feel better now than I ever have before
	I am somewhat ill
	I'm not as healthy now as I used to be
	I'm as healthy as anybody I know
	My health is excellent
	I have been feeling bad lately
	Doctors say that I am now in poor health
	I feel about as good now as I ever have
Prior health	I have been in bed a lot in the past because of illness*
	I was so sick once I thought I might die
	I've never had an illness that lasted a long period of time
	I have never been seriously ill
Resistance/ susceptibility	I seem to get sick a little easier than other people
	Most people get sick a little easier than I do
	My body seems to resist illness very well
	When there is something going around I usually catch it
Health outlook	I think my health will be better in the future than it is now*
	I will probably be sick a lot in the future
	In the near future, I expect to have better health than other people I know
	I expect to have a very healthy life
	Most of the people I know will probably have fewer health problems than I will in the future*
	I think my health will be worse in the future than it is now
Health worry/ concern	I never worry about my health
	I worry about my health more than other people worry about their health
	My health is a concern in my life
	Others seem more concerned about their health than I am about mine
Sickness orientation	Getting sick once in a while is a part of my life
	I accept that sometimes I'm just going to be sick
Rejection of sick role	I try to avoid letting illness interfere with my life
	When I'm sick I try to keep it to myself
	When I'm sick I try to just keep going as usual
	When I think I am getting sick, I fight it
Attitude toward going to the doctor	I don't like to go to the doctor
	It doesn't bother me to go to the doctor

* Item did not satisfy scaling criteria and was not retained.

Using the first method, we factor-analyzed correlations among the eight scales to derive higher-order factors. These factors were interpreted on the basis of the manifest content of the 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 simple orthogonal structure using the methods described previously for item scaling. The higher-order factor structures were nearly identical across field tests; it therefore 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.

On the basis of findings for Form I of the HPQ, it was hypothesized that the prior health and current health scales 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 [1]. It was also hypothesized that the six remaining HPQ scales would correlate highly with the factor(s) defining health outlook and/or sick role propensity and would not correlate highly with the rotated 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 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).

The second method for examining validity was evaluation of relationships among the HPQ scale scores and 11 other measures of health, seven health and illness behaviors, and age. 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 2. In summary, it was hypothesized that favorable health perceptions would be positively associated with positively defined health status variables (e.g., psychological well-being), negatively associated with negatively defined health status variables (e.g., bed days), negatively associated with reporting of illness-related behavior (e.g., number of doctor visits), and would not be associated with health maintenance behavior (check-ups). Relationships were estimated using product-moment correlations and a cross-sectional study design.

It was hypothesized that the two HPQ scales pertaining to sick role propensity (rejection of sick role and attitude toward going to the doctor) 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 and number of doctor visits) because such behaviors reflect both health and the decision to be or not to be sick. Since a high score for attitude toward going to the doctor and a low score for rejection of sick role constitute a propensity for

Table 2. Definition of Validity Variables and Direction of Hypothesized Relation with Health

Variable	Definition	Hypothesized relationship with health
	HEALTH STATUS	
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	−
Psychological well-being	10-item scale to measure positive versus negative psychological well-being	+
Reason for last doctor visit	Whether the last doctor visit was for sickness or injury	−
	HEALTH AND ILLNESS BEHAVIOR	
Number of doctor visits	Number of doctor visits during the prior year	−
Dental visit(s)	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	−

assuming the sick role, it was hypothesized that the two scales would be positively and negatively correlated, respectively, with the behaviors studied.

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 (see Table 1) were eliminated. Rotated and unrotated factor solutions for items and results of discriminant validity tests are presented and discussed in ref. 2. 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., nearly 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.

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) contained only items worded in the same direction.

Distribution of Scale Scores

Means and standard deviations for scores computed from HPQ scales in four field tests are presented in Table 3. The goal of roughly normally dis-

Table 3. Means and Standard Deviations (in parentheses) for HPQ Form II Scales, Four Field Tests

Scale	Scoring range*	Field test			
		A	B	C	D
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 doctor	2–10	5.8 (2.3)	5.8 (2.1)	6.1 (2.4)	6.3 (2.2)

* Lowest possible score is equal to number of items used to compute scale score.

Table 4. Internal-consistency and Test-Retest Reliability Coefficients for Eight HPQ Scales (Four Field Tests)

Scale	No. of items in scale	Internal-consistency coefficient			Test-retest coefficient	
		Low	High	Median	Field test A	Field test B
Current health	9	0.89	0.92	0.91	0.76	0.86
Prior health	3	0.70	0.79	0.73	0.67	0.78
Resistance/susceptibility	4	0.58	0.80	0.71	0.74	0.73
Health outlook	4	0.64	0.79	0.75	0.54	0.76
Health worry/concern	4	0.45	0.62	0.60	0.60	0.65
Sickness orientation	2	0.46	0.61	0.59	0.42	0.72
Rejection of sick role	4	0.54	0.60	0.59	0.41	0.66
Attitude toward going to doctor	2	0.62	0.79	0.67	0.51	0.59

tributed scale scores was nearly achieved. Most of the scale means were somewhat above the midpoints of the scale ranges. 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 field test A (East St. Louis), in which a large proportion of the respondents were disadvantaged.

Reliability of Single-item Scores

Findings regarding the test-retest reliability of HPQ item scores indicated two clear trends. First, item scores tended to be less reliable in field test A. Scores computed for 18 of the 31 items were not sufficiently reliable for purposes of group comparisons in this field test according to the 0.50 standard for reliability coefficients suggested by Helmstadter [36]. Seven of the 31 coefficients did not meet this standard in field test B (Sangamon County). These results may reflect population differences in reliability and/or stability of the health traits measured by these items. Second, coefficients for items in the rejection-of-sick-role group tended to be lowest in both field test A and field test B.

Reliability of Scale Scores

Internal-consistency and test-retest reliability coefficients for HPQ scales in four field tests are summarized in Table 4. Almost without exception, scale scores were sufficiently reliable for purposes of group comparisons (i.e., coefficient $\geqslant 0.50$). In some instances, the current health scale was sufficiently reliable for purposes of individual comparisons, i.e., internal-consistency reliability coefficients $\geqslant 0.90$. 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 and scale measures matched in terms of construct

clearly indicated noteworthy proportional increases in true score variance for scale measures [2]. 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 field test A and from a low of 0.59 to a high of 0.86 in field test B (see Table 4). 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 in which 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 field test A.

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 field test E. 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 [2].

Validity: Higher-order Factor Structure

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, health outlook, health worry/concern, and sickness orientation scales. These results suggest that scales in this group share common variance, presumably due to favorable versus unfavorable perceptions of general health. A second group of scales (rejection of sick role and attitude toward going to the doctor) was unrelated to the unrotated general health factor.

The factors underlying these relationships were further clarified by the rotated higher-order factor structure. Results for four field tests are summarized in Table 5. 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 30 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

Table 5. Percentages of Reliable Variance in Eight HPQ Scales Accounted and Not Accounted for by Higher-order Factors

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

Scale	Reliable variance	Sources of reliable variance*			
		Present/ prior health (factor I)	Future health (factor II)	Sick role propensity (factor III)	Unknown
Current health	91	68	19	1	10
Prior health	74	62	3	0	33
Resistance/susceptibility	70	18	54	6	19
Health outlook	73	26	45	3	20
Health worry/concern	55	5	31	14	43
Sickness orientation	55	2	29	0	65
Rejection of sick role	58	2	2	64	29
Attitude toward going to doctor	69	0	3	54	38

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

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, health outlook, health worry/concern, and sickness orientation) was accounted for by factor II. Although 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 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.

Table 6. Summary of Validity Coefficients for Current Health and Prior Health Scales in Relation to Health Status, Health and Illness Behavior, and Age

Validity variable	Factor I scales			
	Current health		Prior health	
	Proportion significant*	r†	Proportion significant*	r†
HEALTH STATUS				
General health item	1/1	0.80§	1/1	0.50§
Role activity limitations	4/4	–0.46§	4/4	–0.36§
Any bed days	1/1	–0.37§	1/1	–0.21§
Total bed days	1/1	–0.38§	1/1	–0.30§
Sickness	1/1	–0.24§	1/1	–0.18§
Chronic problems	1/1	–0.43§	1/1	–0.32§
Pain	1/1	–0.55§	1/1	–0.39§
Feeling poorly	1/1	–0.28§	1/1	–0.19§
Health worry	1/1	–0.58§	1/1	–0.44§
Psychological well-being	1/1	0.39§	1/1	0.19§
Reason for last doctor visit	1/1	0.21§	1/1	0.11§
HEALTH AND ILLNESS BEHAVIOR				
Number of doctor visits	3/3	–0.42§	3/3	–0.32§
Dental visit(s)	1/3	0.03	2/3	–0.02
Check-up	1/3	–0.05	0/3	–0.03
Doctor visit when needed	1/1	–0.47§	1/1	–0.27§
Recency of care	2/2	0.16§	2/2	0.15§
Hospitalization	3/3	–0.30§	3/3	–0.28§
Compliance	1/1	–0.14§	0/1	–0.07
DEMOGRAPHIC				
Age	4/4	–0.26§	4/4	–0.19§

* Number of significant coefficients in hypothesized direction, in relation to number of field tests.

† Product-moment correlation not corrected for attenuation. In those instances in which more than one correlation was independently computed, table entry is the median across field tests.

§ $P < 0.01$ (one-tailed test).

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 6–8. For this summary three groups of scales were formed on the basis of the higher-order factor structure defined in Table 5.

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 6). Many of these coefficients were moderately high (i.e., above 0.40), indicating a substantial relationship (not corrected for attenuation due to lack of perfect reliability). For both

Table 7. 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	
	Sign.*	r†	Sign.*	r†	Sign.*	r†	Sign.*	r†
HEALTH STATUS								
General health item	1/1	0.50§	1/1	0.58§	1/1	−0.38§	1/1	−0.16§
Role activity limitations	4/4	−0.33§	4/4	−0.24§	4/4	0.19§	4/4	0.14§
Any bed days	1/1	−0.24§	1/1	−0.23§	1/1	0.10§	1/1	0.14§
Total bed days	0/1	0.00	0/1	−0.16	0/1	0.07	0/1	0.01
Sickness	1/1	−0.21§	0/1	−0.13§	1/1	0.12§	1/1	0.12§
Chronic problems	1/1	−0.26§	1/1	−0.34§	1/1	0.10§	1/1	0.14§
Pain	1/1	−0.41§	1/1	−0.35§	1/1	0.24§	1/1	0.18§
Feeling poorly	1/1	−0.21§	1/1	−0.13§	1/1	0.14§	1/1	0.09
Health worry	1/1	−0.42§	1/1	−0.33§	1/1	0.42§	1/1	0.20§
Psychological well-being	1/1	0.18§	1/1	0.28§	1/1	−0.12§	1/1	−0.11§
Reason for last doctor visit	1/1	0.17§	0/1	0.09	0/1	−0.05	0/1	0.01
HEALTH AND ILLNESS BEHAVIOR								
Number of doctor visits	3/3	−0.32§	3/3	−0.16§	3/3	0.20§	1/3	0.10
Dental visit(s)	1/3	−0.02	1/3	0.06	0/3	−0.01	0/3	0.00
Check-up	0/3	−0.01	1/3	−0.01	0/3	−0.04	0/3	0.02
Doctor visit when needed	1/1	−0.21§	1/1	−0.26§	1/1	0.17§	0/1	0.09
Recency of care	2/2	0.15§	1/2	0.06	2/2	−0.15§	0/2	0.00
Hospitalization	3/3	−0.26§	2/3	−0.11§	1/3	−0.07	0/3	0.02
Compliance	0/1	−0.06	0/1	−0.06	1/1	0.16§	0/1	−0.10
DEMOGRAPHIC								
Age	1/4	0.09	4/4	−0.23§	3/4	−0.10§	2/4	−0.06

* Number of significant coefficients in the hypothesized direction, in relation to number of field tests.

† Product-moment correlation not corrected for attenuation. In those instances in which more than one correlation was independently computed, table entry is the median across field tests.

§ $P < 0.01$ (one-tailed test).

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 item, 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 (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.

Table 8. 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	
	Sign.*	r†	Sign.*	r†
HEALTH STATUS				
General health item	0/1	0.02	1/1	0.23§
Role activity limitations	0/4	–0.03	0/4	–0.03
Any bed days	0/1	0.01	0/1	–0.03
Total bed days	0/1	0.06	1/1	0.22§
Sickness	0/1	–0.04	0/1	–0.03
Chronic problem(s)	0/1	–0.08	0/1	0.00
Pain	1/1	–0.11§	1/1	–0.13§
Feeling poorly	0/1	–0.03	0/1	–0.07
Health worry	1/1	–0.12§	1/1	–0.13§
Psychological well-being	0/1	0.04	0/1	0.05
Reason for last doctor visit	0/1	0.00	0/1	0.07
HEALTH AND ILLNESS BEHAVIOR				
Number of doctor visits	0/3	–0.02	1/3	0.08
Dental visit(s)	0/3	–0.01	1/3	0.09
Check-up	0/3	0.00	2/3	0.15§
Doctor visit when needed	0/1	0.01	0/1	0.00
Recency of care	0/2	–0.03	2/2	–0.11§
Hospitalization	1/3	–0.10	0/3	0.00
Compliance	0/1	–0.03	1/1	0.21§
DEMOGRAPHIC				
Age	4/4	0.14§	0/4	0.04

* Number of significant coefficients in the hypothesized direction, in relation to number of field tests.

† Product-moment correlation not corrected for attenuation. In those instances in which more than one correlation was independently computed, table entry is the median across field tests.

§ $P < 0.01$ (one-tailed test).

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 7). Correlations tended to be more consistently significant and higher in absolute magnitude for health outlook and resistance/susceptibility 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 who reported one or more). It is interesting that the latter pattern was not observed for the two factor I scales.

Resistance/susceptibility, 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 with medical regimens (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 8); correlations that were significant (e.g., with pain and 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. Significant correlations between factor III scales and behaviorally defined health variables were as hypothesized; the correlation with the general health rating item was not. The absence of consistent and substantial relationships between factor III scales and the other health variables studied is consistent with the hypothesis that these 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 one of three field tests).

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

Discussion

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 clearly measure the same primary constructs. The fact that item groupings as hypothesized were empirically validated in 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.

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 most studies of general health ratings published in the last 25 years have been based on single-item measures [4] 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 it takes only about seven minutes for the average respondent to answer the 32 HPQ items necessary to construct the eight scales. When reliability coefficients for disadvantaged and nondisadvantaged respondents were compared, differences between test-retest coefficients were considerably greater than differences between internal-consistency coefficients. These results 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 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 [12,15,20]; 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

Relationships Among HPQ Scales. Results of the empirical studies of relationships among the eight HPQ scales clearly indicate that each scale score primarily reflects one of three kinds of perceptions: 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 other constructs. The findings constitute strong support for construct validity because the results were very clear for most scales and because the results of independent field tests were nearly identical.

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 use of the scales. In this regard it is recommended that Table 5 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, 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 trade-offs.

Other Health Variables. Findings regarding the relationships among the eight scales and higher-order factors pertaining to health and sick role propensity constitute support for the hypothesis 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 the other two scales would not. 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 (rejection of sick role and attitude toward going to the doctor) 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.

Although 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 [2], 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. Given that prior behaviors are certain to have affected present ratings [2,39], and given the effects of methodological problems [40], the methods used in the present study 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 [9,19,23]. Only current health and health worry/concern were significantly related to compliance with prescribed medical regimens. Consistent with previous findings [1,23] and current study hypotheses, health perceptions were not related to health behavior (e.g., check-ups). The fact 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) may have been an artifact of the cross-sectional study design. A test of the validity of the HPQ scales with a prospective study design is under way [41].

The results obtained for attitude toward going to the doctor and rejection of the sick role scales 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).

Since correlations among HPQ scales were not taken into account in

evaluating their validities in relation to behavior, it was not possible to determine from these results whether each of the scales makes a unique contribution to an explanatory model. This issue is addressed elsewhere [2].

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 [1,11,13,24,25]. When examined for nonlinearity, relationships between age and general health perceptions, when significant, tended to be linear [2]. 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 [1].

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. Although further research is necessary to understand the clinical significance of scale scores, the information obtained 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 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 development and validation of shorter forms of the scales, instruments for use with children, and on the validity of the scales in prospective study designs, which will be necessary to clarify the causes and effects that are operating in relationships among these and other health constructs as well as in health and illness behavior.

Acknowledgments. I gratefully acknowledge 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 H. 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 through 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.

References

1. 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.* PB-242-726. Springfield, VA: National Technical Information Service, 1974.
2. Ware, J.E. Jr. and A.H. Karmos. *Development and Validation of Scales to Measure Perceived Health and Patient Role Propensity.* School of Medicine, Southern Illinois University, Carbondale, IL, 1976.
3. Ware, J.E. Jr., M.K. Snyder, E.R. McClure, and I.W. Jarret. *The Measurement of Health Concepts.* PB-239-508/AS. Springfield, VA: National Technical Information Service, Dec. 1972.
4. Donald, C.A. and J.E. Ware Jr. A review of the literature on general health ratings. Unpublished paper, Rand Corp., Santa Monica, CA, 1976.
5. 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 Mar. 1971.
6. Hochstim, J.R., D.A. Athanasopoulos, and J.H. Larkins. Poverty area under the microscope. *Am J Public Health* 58:1815 Oct. 1968.
7. Madow, W.G. *Interview Data on Chronic Conditions: Compared with Information from Medical Records.* PHS Pub. No. 1000. Series 2, No. 23. Washington, DC: U.S. Government Printing Office, 1967.
8. Edwards, J. and D. Klemmack. Correlates of life satisfaction: A reexamination. *J Gerontol* 28:497 Oct. 1973.
9. 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.
10. Reynolds, J., W.A. Rushing, and D.L. Miles. The validation of a Function Status Index. *J Health Soc Behav* 15:271 Dec. 1974.
11. Schnore, L.F. and J.D. Cowhig. Some correlates of reported health in metropolitan centers. *Soc Probl* 7:218 1959–60.
12. Suchman, E.A., B.S. Phillips, and G.F. Streib. An analysis of the validity of health questionnaires. *Soc Forces* 26:223 Mar. 1958.
13. Cole, S. and R. Lejeune. Illness and the legitimation of failure. *Am Sociol Rev* 37:347 June 1972.
14. Cutler, S. Voluntary association participation and life satisfaction: A cautionary research note. *J Gerontol* 28:96 Jan. 1973.
15. Maddox, G.L. Self-assessment of health status: A longitudinal study of selected elderly subjects. *J Chron Dis* 17:449 May 1964.
16. Renne, K.S. Correlates of dissatisfaction in marriage. *J Marriage Fam* 22:54 Feb. 1970.
17. Renne, K.S. Health and marital experience in an urban population. *J Marriage Fam* 23:338 May 1971.
18. Renne, K.S. Measurement of social health in a general population survey. *Soc Sci Res* 3:25 Mar. 1974.
19. Friedsam, H.J. and H.W. Martin. A comparison of self and physicians' health ratings in an older population. *J Health Hum Behav* 4:179 Fall 1963.
20. 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 Apr. 1963.
21. Kessel, N. and M. Shepard. The health and attitudes of people who seldom consult a doctor. *Med Care* 3:6 Jan.–Mar. 1965.
22. Ludwig, E.G. and G. Gibson. Self-perception of sickness and the seeking of medical care. *J Health Soc Behav* 10:125 June 1969.
23. Tissue, T. Another look at self-rated health among the elderly. *J Gerontol* 27:91 Jan. 1972.
24. Gaitz, C.M. and J. Scott. Age and the measurement of mental health. *J Health Soc Behav* 12:55 Mar. 1972.
25. Palmore, E. and C. Luikart. Health and social factors related to life satisfaction. *J Health Soc Behav* 13:68 Mar. 1972.
26. Spreitzer, E. and E.E. Snyder. Correlates of life satisfaction among the aged. *J Gerontol* 29:454 July 1974.
27. Pratt, L. The relationship of socioeconomic status to health. *Am J Public Health* 61: 281 Feb. 1971.

28. Macpherson, A.S. *The Measurement of Mental Health in a General Population.* Master's thesis, School of Graduate Studies, McMaster University, Hamilton, Ontario, 1972.
29. Stevens, S.S. Mathematics, Measurement, and Psychophysics. In S.S. Stevens (ed.), *Handbook of Experimental Psychology,* pp. 1–49. New York: Wiley, 1951.
30. Ware, J.E. Jr., W.R. Wright, and M.K. Snyder. *Development and Validation of Scales to Measure Key Health Concepts: Methodological Appendix.* Southern Illinois University, Carbondale, IL, 1976.
31. Comrey, A.L. *A First Course in Factor Analysis.* New York: Academic Press, 1973.
32. Comrey, A.L. Tandem criteria for analytic rotation in factor analysis. *Psychometrika* 32:143 June 1967.
33. Campbell, D.T. and F.W. Fiske. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull* 56:81 Mar. 1959.
34. Howard, K.I. and G.G. Forehand. A method for correcting item-total correlations for the effect of relevant item inclusion. *Educ Psychol Meas* 22:731 Winter 1962.
35. Cronbach, L.J. Coefficient Alpha and the internal structure of tests. *Psychometrika* 16:297 Sept. 1951.
36. Helmstadter, G.C. *Principles of Psychological Measurement,* pp. 355–61. New York: Appleton-Century-Crofts, 1973.
37. Stewart, A., J.E. Ware Jr., and R.H. Brook. The meaning of health: Understanding functional limitations. *Med Care* (in press).
38. Brook, R.H. and J.E. Ware Jr. *Conceptualization and Measurement of Health for Adults in the Health Insurance Study: Dayton.* R-1987-HEW. Santa Monica, CA: Rand (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.* Southern Illinois University, Carbondale, IL, 1976.
40. Ware, J.E. Jr. *A Cross-sectional Study of Patient Satisfaction and Use of Health Care Services in Rural Southern Illinois.* P-5691. Santa Monica, CA: Rand (forthcoming).
41. Newhouse, J.P. A design for a health insurance experiment. *Inquiry* 11:5 Mar. 1974.