



H • S • I

RAND-36 Health Status Inventory

Ron D. Hays

With special contributions by
Sandra Prince-Embury and Hsin-Yi Chen



THE PSYCHOLOGICAL CORPORATION®
Harcourt Brace & Company

SAN ANTONIO

Orlando • Boston • New York • Chicago • San Francisco • Atlanta • Dallas
San Diego • Philadelphia • Austin • Fort Worth • Toronto • London • Sydney

Copyright ©1998 by The Psychological Corporation

Normative data copyright ©1998 by The Psychological Corporation

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

The Psychological Corporation and the *PSI* logo are registered trademarks of The Psychological Corporation.

Beck Depression Inventory, BDI, Beck Anxiety Inventory, BAI, Beck Hopelessness Scale, and BHS are registered trademarks of The Psychological Corporation.

Printed in the United States of America

123456789101112ABCDE

Foreword

The 36 questions that comprise the *RAND-36 Health Status Inventory* are probably the most commonly asked health status questions worldwide. Also known as the SF-36, this questionnaire is being used in countless health outcome studies, health care financing studies, and clinical practice evaluations. Although there is an increasing volume of published data from various diagnostic cohorts and defined populations, including the original Medical Outcomes Study (MOS) standardization sample, there has not to date been a manual that provides U.S. census-based norms stratified by sex, age, racial/ethnic group, and educational level. These data could well become the benchmark for comparison of one's local, regional, or national results.

The 36 questions in the RAND-36 HSI were selected from the larger pool of items used in the MOS. Item selection was based on criteria that maximized item association with the longer scales. These 36 items were also selected in part to ensure coverage of the full spectrum of physical and mental health. One consequence of the commitment to cover a full range of functioning with relatively few items is the loss of precision and sensitivity to change at any given level of health, compared to disease or condition-specific assessment. However, a decided gain of this approach is the distribution of item difficulty across the continuum of health, as measured by item response theory (IRT). Therefore, although this questionnaire was not developed by IRT methodology, it is well suited to it. This Manual represents the first major effort to produce a scoring system for these 36 items that capitalizes on the strengths of IRT. The IRT method of ordering all items, and response categories within items, along a single continuum allows for the empirical weighting of responses to questions of differing difficulty according to that underlying continuum. This approach is a major advance in scoring of responses to the questionnaire and can, if developed, evolve into a major advance in health status assessment.

This Manual provides an unparalleled matrix of normative data for group and individual comparison purposes. The use of nonorthogonal factor rotations to derive separate physical and mental health composite scores is sensible, given the known relationship between these components of health, and distinguishes the RAND-36 HSI from the SF-36 scoring system. In addition, the use of a Global Health Composite score is unique and (I daresay) long awaited by the scientific community. Several health status questionnaires that tap physical and mental health domains concurrently have managed to create workable, psychometrically sound total scores. In our work with cancer and HIV patients, we have computed an IRT-based total score from the RAND-36 HSI, but this Manual marks the first time that a formal scoring system will be disseminated along with population-based normative data for adults. The Global Health Composite score has somehow previously eluded formal scoring systems for such questionnaires, but it is successfully implemented here with use of IRT methodology. This total score will have value not only at the group decision-making level but also at the individual patient-management level, where physical and mental health concerns must frequently be balanced with each other and combined into a single "bottom-line" summary.

This Manual includes over 50 tables, most of which provide *T*-score conversions for each of the individual scale and composite scores. These conversions are helpful because they transform IRT-based logistic data, which might be unfamiliar for some users, into a more familiar, standardized metric. These tables will no doubt be helpful to the benchmarking of an individual user's data, but perhaps the most useful aspects of this Manual are the sections on determining statistical significance of change and evaluating the clinical meaningfulness of change scores. Without this kind of information, interpretation of results is limited. Information about the clinical, real-world relevance of a given score or an increment of change in that score is vital to interpretation. This information, in turn, may move health care providers toward better cost-effectiveness and cost-utility modeling as we understand just what the meaning of an improvement in health afforded by a given intervention is.

In conclusion, there are thousands of people using these 36 questions in their clinical practice and clinical research. Most will benefit from having this Manual within reach when making sense of their data. Perhaps this work, carried out so meticulously with the U.S. English-speaking population, will next be expanded to include other languages and other countries where these same 36 questions are used.

David Cella, PhD
Research Professor
Institute for Health Services Research and Policy Studies
Northwestern University
Director, Center on Outcomes Research and Education
Evanston Northwestern Healthcare

Acknowledgments

I would like to acknowledge a number of people for their outstanding contributions to the development of the *RAND-36 Health Status Inventory* (RAND-36 HSI). Sandra Prince-Embury, PhD, Project Director, did an amazing job from start to finish in seeing this effort through to completion. Dr. Prince-Embury is really the major author of this work and deserves most of the credit for its completion. Hsin-Yi Chen, PhD, was the primary architect for the application of item response theory scaling in this project and played a central role in most of the statistical decisions. Larry Weiss, PhD, Director of Behavioral Healthcare and Personality Group, provided valuable consultation throughout the project, including the design of the data-collection plan and help in preparing this Manual and the normative tables. Appreciation also goes to David Tulsy, PhD, Project Director, for reviewing the Manual. Stephanie Tong, Research Assistant, dedicated many hours to the preparation of tables and revisions of the Manual. Her diligence and attention to detail in ensuring that the information presented is accurate and complete are especially valued.

Special thanks are also extended to those individuals whose meticulous and diligent efforts were essential in the preparation of the Manual for publication. Those responsible for ensuring the editorial quality of the Manual are Kathy Overstreet, Senior Editor, and Cynthia Woerner, Consulting Editor. The production of the Manual itself was in the capable hands of Debra Belknap, Senior Production Manager; Margaret Donohue, Production Manager; Michael Friedman, Production Specialist; Pat Malec, Manager of Art and Design; and Javier Flores, Associate Designer.

Joanne Lenke, PhD, President of The Psychological Corporation, and Aurelio Prifitera, PhD, Vice President and Director of the Psychological Measurement Group, provided administrative support for the project. Dr. Prifitera also personally reviewed the Manual and provided helpful feedback.

A special thanks is due to David Cella, PhD, a pioneer in using item response theory methods in health status assessment, for his diligence in reviewing the Manual and writing the foreword.

Finally, this effort would not have been possible without the contributions of the measurement experts at the RAND Corporation who worked on the RAND Medical Outcomes Study (MOS). These individuals include Anita Stewart, Cathy Sherbourne, and the principal investigator of the MOS, John Ware.

Ron D. Hays, PhD
RAND Health Sciences Program, Santa Monica, California
UCLA School of Medicine, Los Angeles, California
July 1998

Contents

Foreword	iii
Acknowledgments	v
Chapter 1 Introduction	
Constructs Assessed	2
Scales	2
Physical Functioning	2
Role Limitations due to Physical Health Problems	3
Pain	3
General Health Perceptions	3
Emotional Well-Being	3
Role Limitations due to Emotional Problems	3
Social Functioning	3
Energy/Fatigue	3
Composites	3
Reliability and Validity	4
Research and Clinical Applications	5
Target Populations	5
Individual-Level Assessment	5
Chapter 2 Standardization	
Description of the Sampling Procedure	7
Demographic Characteristics of the Age-Based Sample	8
Demographic Characteristics of the Age-Stratified Sample	11
Disability Status and Physical Conditions	12
Chapter 3 Item Scaling and Development of the Composites	
Item Scaling	15
Theoretical Background	15
Previous Research	16
Application of IRT Weighting	17

Effects of IRT Scoring on Distribution	17
Development of the Composites	18

Chapter 4 Norming

Determination of Normative Groups	21
Standardization of Scale Raw Scores	24
Standardization of Composite Raw Scores	24
Cumulative Percentages for <i>T</i> Scores	24

Chapter 5 Reliability and Validity of Scores

Reliability	25
Reliability of Scale Scores	25
Reliability of Composite Scores	26
Test-Retest Reliability	26
Validity	27
Item Level	27
Scale Level	28
Composite Level	28
Previous Studies	29
Association of RAND-36 HSI With Other Measures	29
Physical Health Measures	30
Mental Health Measures	31
Global Health Composite	34
Known-Group Analyses	34
Utilization of Health-Care Resources	35

Chapter 6 Administration and Scoring

Hand-Scoring for the RAND-36 HSI	39
Hand-Scoring for the RAND-12 HSI	39
Computer-Based Administration, Scoring, and Interpretation	40

Chapter 7 Interpretation of the RAND-36 HSI Scores

Determination of Clinical Meaningfulness	46
General Parameters	46
Missing Item Responses	46
Clinical Cut Scores	46

Sequential Analysis	48
Suggested Guidelines	48
Physical Health Composite	48
Mental Health Composite	48
Physical Health and Mental Health Composite Score Discrepancy ...	49
Global Health Composite	49
Scale Scores	50
A Final Caution	50
Longitudinal Tracking of Change	50
General Guidelines	51
Step 1: Determining Statistical Significance of Change	52
Step 2: Evaluating the Clinical Meaningfulness of Change	53
Summary	54

Appendixes

Appendix A Computation of Item Scores	55
Appendix B Computation of Scale and Composite Scores for the RAND-36 HSI	61
Appendix C <i>T</i> Scores Obtained by Cumulative Percentages of the Normative Samples	85
Appendix D Discrepancy Score Tables	105
Appendix E Description of the <i>RAND-12 Health Status Inventory</i>	107
Appendix F RAND-36 HSI Items and Response Options by Composite and Scale	115
Appendix G RAND-12 HSI Items and Response Options by Composite ..	119

References

Figures

Figure 2.1.	Percentages of the RAND-36 HSI Age-Based Standardization Sample and the U.S. Population by Occupation	11
Figure 6.1.	OPTAIO Session Graph of RAND-36 HSI Composite and Scale Scores	40
Figure 6.2.	OPTAIO Longitudinal Graph of RAND-36 HSI Composite Scores	41
Figure 6.3A.	OPTAIO Interpretive Report of RAND-36 HSI Performance (Page 1)	42
Figure 6.3B.	OPTAIO Interpretive Report of RAND-36 HSI Performance (Page 2)	43

Tables

Table 2.1.	Percentages of the RAND-36 HSI Age-Based Standardization Sample and the U.S. Population by Race/Ethnicity	9
Table 2.2.	Percentages of the RAND-36 HSI Age-Based Standardization Sample and the U.S. Population by Education Level	9
Table 2.3.	Median Ages of the RAND-36 HSI Age-Based Standardization Sample by Age Group and Sex	10
Table 2.4.	Percentages of Unemployed Participants in the RAND-36 HSI Age-Based Standardization Sample and the U.S. Population by Sex	10
Table 2.5.	Percentages of the RAND-36 HSI Age-Stratified Sample and the U.S. Population by Age Group and Sex	12
Table 2.6.	Percentages of the RAND-36 HSI Age-Stratified Sample and the U.S. Population by Race/Ethnicity and Sex	12
Table 2.7.	Percentages of the RAND-36 HSI Age-Stratified Sample and the U.S. Population by Education Level and Sex	12
Table 2.8.	Percentages of RAND-36 HSI Participants Reporting Disabilities	13
Table 2.9.	Percentages of Physical Conditions and Mean Level of Interference in Functioning Reported by RAND-36 HSI Standardization Participants	14
Table 3.1.	Promax Factor Pattern Loadings for RAND-36 HSI Scales	20
Table 4.1.	Mean Scale Raw Scores, Standard Deviations, Range Scores, and Distribution Statistics for the Seven Normative Groups	22
Table 5.1.	Reliability Coefficients for RAND-36 HSI Scale and Composite Scores for the Seven Normative Groups	26

Table 5.2.	Test-Retest Reliability Coefficients for RAND-36 HSI Scale and Composite Scores	27
Table 5.3.	Intercorrelations of the RAND-36 HSI Scale Scores for the Age-Stratified Sample	28
Table 5.4.	Promax Factor Pattern Loadings for the RAND-36 HSI Scales	29
Table 5.5.	Correlations Between the RAND-36 HSI and Other Indicators of Physical Health	30
Table 5.6.	Correlations Between the RAND-36 HSI and the BDI-II, BAI, and BHS for the Age-Stratified Sample	31
Table 5.7.	Correlations Between the RAND-36 HSI and the <i>Brief Symptom Inventory</i>	32
Table 5.8.	Correlations Between the RAND-36 HSI and the BASIS-32	33
Table 5.9.	Correlations Between the RAND-36 HSI and the SAS-SR	34
Table 5.10.	Correlations Between the RAND-36 HSI and Self-Reported Health-Care Resource Utilization	37
Table 7.1.	Standard Errors of Prediction for the Three RAND-36 HSI Composites at 90% Level of Confidence	53
Table 7.2.	<i>T</i> -Score Ranges for Evaluating the Clinical Meaningfulness of Change	54
Table A.1.	Item Response-Option Weights: Physical Functioning Scale (PF)	56
Table A.2.	Item Response-Option Weights: Role Limitations due to Physical Health Problems Scale (RLP)	56
Table A.3.	Item Response-Option Weights: Pain Scale (PA)	57
Table A.4.	Item Response-Option Weights: General Health Perceptions Scale (GHP)	57
Table A.5.	Item Response-Option Weights: Emotional Well-Being Scale (EWB)	58
Table A.6.	Item Response-Option Weights: Role Limitations due to Emotional Problems Scale (RLE)	58
Table A.7.	Item Response-Option Weights: Social Functioning Scale (SF)	59
Table A.8.	Item Response-Option Weights: Energy/Fatigue Scale (EF)	59
Table B.1.	Raw-Score Ranges for the Scales	63
Table B.2.	<i>T</i> -Score Equivalents of Raw Scores: Physical Functioning Scale (PF)	64

Table B.3.	<i>T</i> -Score Equivalents of Raw Scores: Role Limitations due to Physical Health Problems Scale (RLP)	66
Table B.4.	<i>T</i> -Score Equivalents of Raw Scores: Pain Scale (PA)	68
Table B.5.	<i>T</i> -Score Equivalents of Raw Scores: General Health Perceptions Scale (GHP)	70
Table B.6.	<i>T</i> -Score Equivalents of Raw Scores: Emotional Well-Being Scale (EWB)	72
Table B.7.	<i>T</i> -Score Equivalents of Raw Scores: Role Limitations due to Emotional Problems Scale (RLE)	74
Table B.8.	<i>T</i> -Score Equivalents of Raw Scores: Social Functioning Scale (SF)	75
Table B.9.	<i>T</i> -Score Equivalents of Raw Scores: Energy/Fatigue Scale (EF)	77
Table B.10.	<i>T</i> -Score Equivalents of Raw Scores: Physical Health Composite (PHC)	79
Table B.11.	<i>T</i> -Score Equivalents of Raw Scores: Mental Health Composite (MHC)	81
Table B.12.	<i>T</i> -Score Equivalents of Raw Scores: Global Health Composite (GHC)	83
Table C.1.	<i>T</i> Scores Obtained by Cumulative Percentages of the Normative Samples: Physical Functioning Scale (PF)	85
Table C.2.	<i>T</i> Scores Obtained by Cumulative Percentages of the Normative Samples: Role Limitations due to Physical Health Problems Scale (RLP)	87
Table C.3.	<i>T</i> Scores Obtained by Cumulative Percentages of the Normative Samples: Pain Scale (PA)	88
Table C.4.	<i>T</i> Scores Obtained by Cumulative Percentages of the Normative Samples: General Health Perceptions Scale (GHP)	90
Table C.5.	<i>T</i> Scores Obtained by Cumulative Percentages of the Normative Samples: Emotional Well-Being Scale (EWB)	92
Table C.6.	<i>T</i> Scores Obtained by Cumulative Percentages of the Normative Samples: Role Limitations due to Emotional Problems Scale (RLE)	94
Table C.7.	<i>T</i> Scores Obtained by Cumulative Percentages of the Normative Samples: Social Functioning Scale (SF)	95
Table C.8.	<i>T</i> Scores Obtained by Cumulative Percentages of the Normative Samples: Energy/Fatigue Scale (EF)	97
Table C.9.	<i>T</i> Scores Obtained by Cumulative Percentages of the Normative Samples: Physical Health Composite (PHC)	99

Table C.10.	<i>T</i> Scores Obtained by Cumulative Percentages of the Normative Samples: Mental Health Composite (MHC)	101
Table C.11.	<i>T</i> Scores Obtained by Cumulative Percentages of the Normative Samples: Global Health Composite (GHC)	103
Table D.1.	Cumulative Percentages of the Normative Samples Obtaining PHC>MHC <i>T</i> -Score Discrepancies	105
Table D.2.	Cumulative Percentages of the Normative Samples Obtaining MHC>PHC <i>T</i> -Score Discrepancies	106
Table E.1.	Correlations Between RAND-12 HSI and RAND-36 HSI Item Scores for the Seven Normative Groups	112
Table E.2.	Intercept Values and Parameter Estimates for Predicting Physical Health Composite <i>T</i> Scores From the RAND-12 HSI	113
Table E.3.	Intercept Values and Parameter Estimates for Predicting Mental Health Composite <i>T</i> Scores From the RAND-12 HSI	113
Table E.4.	Intercept Values and Parameter Estimates for Predicting Global Health Composite <i>T</i> Scores From the RAND-12 HSI	114
Table E.5.	Cross-Validation Results Predicting RAND-36 HSI Composite <i>T</i> Scores From the RAND-12 HSI	114