Wright's Pack - Team Theory

RASCH MODELS FOR N-MEMBER GROUPS

PACK: $B_{\Sigma n} - D \approx \Sigma(B_n - D)/N + \log(N)$

TEAM: $B_{\Pi n} - D = \Sigma(B_n - D)$

RASCH MEASUREMENT

Transactions of the Rasch Measurement SIG
American Educational Research Association

Vol. 22 No. 4 | Spring 2009 | ISSN 1051-0796

Your Invitation to "A Tribute to Benjamin D. Wright"

Please join us for a tribute to Benjamin D. Wright

When: Saturday, March 28, 2009
6:00 PM

Where: the Wyndham Chicago
633 N Saint Clair St
Chicago, Ilinois 60611

Details: No-host dinner
Cost: $45 per person

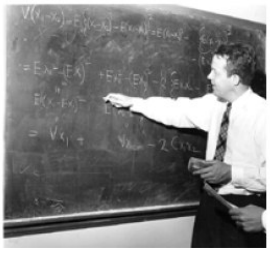
RSVP: By March 18, 2009

Please send a check payable to the Rehabilitation Instiute of Chicago to:

Holly DeMark
Rehabilitation Institute of Chicago
345 E. Superior St. / CROR
Chicago, IL 60611

Questions? Contact Holly at (312) 238-2826 • hdemark@ric.org

Benjamin D. Wright celebrates his 83rd birthday on March 30. The measurement community celebrates his lifetime of achievements on *Saturday, March 28, 2009.*



Dr. Wright has published more than 150 papers on Rasch measurement,co-authored 12 books, and directed the development of the two most widely used Rasch measurement computer programs. His leadership in the field is unparalleled, and many of his students are contemporary leaders in psychometrics.

The evening will consist of informal speeches and a photo slide show to commemorate Dr. Wright. Moderators and speakers include Ed Bouchard, Allen Heinemann, Nikolaus Bezrucko, and many others.

Ben Wright will speak at the celebratory Symposium: *"Improving Efficiency in Health Outcome Measurement"*, Chicago, Saturday, March 28, 2009.

www.ric.org/research/centers/cror/projects/rrtc/data/T3.aspx

Table of Contents

Indexing vs. Measuring

In RMT 22:1, Stenner, Stone, and Burdick (2008) distinguished between two different measurement models: reflective or latent variable models and formative or composite variable models (Edwards & Bagozzi, 2000). In the former, the causal action flows from the latent variable to the indicators (e.g., temperature) whereas, in the latter, the causal action flows from indicators to the composite variable (e.g., socioeconomic status). We believe that the language we use should accentuate these differences and as such we propose to call reflective models *measurement models,* what these models produce we will call *measures* and the process of producing these measures will be called measuring. In parallel fashion, formative models will be called *index models,* what they produce we will call indices, and the process of producing *indices* will be called *indexing.* The notion of an index is well developed in economics and sociology and carries the connotations we desire. What follows is a discussion of how indexing and measuring differ and why it is important to make this distinction in the human sciences.

Definitions

***Indices* are the effects of their indicators whereas *measures* (of latent variables) are the causes of their indicators.** So, changes in stature or consumer price behavior are caused by changes in height (or weight) and price changes for market baskets of commodities (computers, milk, gasoline), respectively. Changes in latent variable measures, in contrast, cause a homogeneous (often nonlinear) change in indicator behavior, as when temperature change causes thermometric fluid to expand in the thermometer or a change in reader ability causes a change in count correct on a reading test.

Altering the indicators of an index changes the definition of the variable being indexed, whereas changing the indicators for a measure will not alter the latent variable (although precision of measurement and or unit size may be affected). So, if midline girth is added to height and weight as indicators of stature or all electronic commodities are eliminated from the Consumer-Product-Index (CPI) market basket, the definition of what is being indexed changes.

In contrast, knowledge of expansion coefficients and viscosity differences allows us to swap new thermometric fluids for mercury without changing the construct being measured. Similarly, new reading items with different text and item types can be swapped for previous items without changing the construct being measured.

Another way to express this point is that the indicators for an index are constitutive of that index, whereas indicators for a latent variable are incidental to the construct's definition.

Specification Equations

In a generally objective measurement framework (e.g., Lexiles) what is crucial in the definition of the construct is the specification equation that specifies the cause of the variation detected by the instrument. Because the indicators of an index by design track different kinds of variation (height, weight, midline girth), it is difficult to imagine a specification equation that could, somehow, capture what these indicators share independent of the linear (or otherwise) combination that constitutes the index. What, for example, would a parallel form of Sheldon's *somatotype* rating scale look like? Difficulty in imagining what new indicators would constitute a parallel form is strongly suggestive of the need for an index rather than a measurement model.

Indices Misinterpreted as Latent Variables

Because both index and measurement models are fundamentally associational (i.e., based on correlations among indicators), traditional applications of Rasch model software often cannot distinguish between an index and a latent variable (Stenner, Burdick, & Stone, 2008). Examples of resulting confusion take predominantly one particular form: index variables are interpreted, as if they are latent variables. Here is an example typical of many in the Rasch literature *[and RMT, Ed.]:*

The Rasch model has been shown to fit FIM data reasonably well, which indicates that the scale locations describe adequately the relative order in which these functions are lost in the aging population. The items on the top describe difficult activities, such as climbing stairs, whereas items on the bottom describe easier activities that are maintained relatively well. (Embretson, 2006, p. 52)

Contrary to a latent variable interpretation the FIM (Functional Independence Measure) appears to be an index of motor functioning with the causal action moving from indicators to index. If the desired medical outcome is "more functional independence," then rehabilitating bladder control, walking, bathing, and so on should promote the intended outcome rather than the other way around. Alternatively, we could teach the patient to drive a motorized wheelchair but to include this as an indicator would alter the definition of functional independence.

Global fit of data to a Rasch model will not sort out the direction of causal flow and thus will not provide unambiguous evidence for a latent variable interpretation of the construct. A substantive theory and associated specification equation capable of explaining variation in indicator difficulties is a big step in support of a latent variable interpretation. The coup de grace is a demonstration of the specification equation's causal status using experimental manipulation of instrument characteristics (radicals) and subsequent observation of

the theoretically predicted change in the measurement outcome.

Correlation is not Causation

It is a property of indices (economical, sociological, or psychological) that the indicator composite may be found to correlate more highly with an unintended criterion than the intended one. Such a discomforting outcome is yet another reason that a correlational (as opposed to a causal) view of validity is not sustainable.

Latent variable interpretations are most defensible when global fit of data to a Rasch model is accompanied by invariance of the indicator structure throughout the range of the construct. In the language of additive conjoint measurement (Luce & Tukey, 1964) and as realized in the Lexile Framework for reading (Kingdon in press), it should be possible to trade off a difference between reader abilities of 200L for a difference in text readability of 200L to hold comprehension rate (count correct/total items) constant (Burdick, Stone, & Stenner, 2006). This trade-off property has been shown to operate throughout the grade range from kindergarten to advanced adult reading (e.g., Supreme Court decisions) and would not be expected to hold for a reading index variable composed of items such as: (1) number of books in the home, (2) daily newspaper subscription, (3) English as a first language, etc.

It may be true that "where there is correlational smoke there is likely to be causational fire" (Holland, 1986, p. 951). Good fit with a Rasch model is correlational smoke, but as we have just seen, it takes an experimental test of a substantive theory to unambiguously distinguish between a latent variable and an index.

A. Jackson Stenner, Mark H. Stone, and Donald S. Burdick

Burdick, D. S., Stone, M. H., & Stenner, A. J. (2006). The Combined Gas Law and a Rasch Reading Law. Rasch Measurement Transactions, 20(2), 1059-60, www.rasch.org/rmt/rmt202.pdf

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. Psychological Methods, 5, 155-174.

Embretson, S. E. (2006). The continued search for nonarbitrary metrics in psychology. American Psychologist, 61(1), 50-55.

Holland, P. W. (1986). Statistics and causal inference. Journal of the American Statistical Association, 81, 945-960.

Luce, R & Tukey, J. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. Journal of Mathematical Psychology, 1, 1-27.

Stenner, A. J., Burdick, D. S., & Stone, M. H. (2008). Formative and reflective models: Can a Rasch analysis tell the difference? Rasch Measurement Transactions, 22:1, 1152-3 www.rasch.org/rmt/rmt221.pdf

Rasch-related Coming Events

March 15, 2009, Sun. Online Programs in Research Methodology (Fall 2009, registration deadline), www.rasch.org/onlineuic.htm

March 23-24, 2009, Mon.-Tues. Introduction to Rasch Measurement and Winsteps (N. Lide), Malaysia www.iiu.edu.my/visitor/detail.shtml?eventid=564

March 28, 2009, Sat. Symposium: Improving Efficiency in Health Outcome Measurement, Chicago www.ric.org/research/centers/cror/projects/rrtc/data/T3.aspx

April 13-17, 2009, Mon.-Fri. AERA Annual Meeting, San Diego, CA, USA, www.aera.net

May 1-29, 2009, Fri.-Fri. Many-Facet Rasch Measurement online course (M. Linacre, Facets), www.statistics.com/ourcourses/facets

May 13-15, 2009, Wed.-Fri. Introduction to Rasch (A. Tennant, RUMM), Leeds, UK, www.leeds.ac.uk/medicine/rehabmed/psychometric

May 18-20, 2009, Mon.-Wed. Intermediate Rasch (A. Tennant, RUMM), Leeds, UK, www.leeds.ac.uk/medicine/rehabmed/psychometric

June 2-3, 2009, Tues.-Wed. 2009 GMAC Invitational Conference on Computerized Adaptive Testing, Minneapolis, www.gmac.com/CATConference

June 26 - July 24, 2009, Fri.-Fri. Rasch - Core Topics online course (M. Linacre, Winsteps), www.statistics.com/ourcourses/rasch1

June 29 - July 2, 2009, Mon.-Thur. 9th International Symposium on Measurement Technology and Intelligent Instruments, Russia, www.tdisie.nsc.ru/ismtii2009

July 20 - Nov. 14, 2009, Mon.-Sat. Introduction to Rasch Measurement of Modern Test Theory online course (D. Andrich, I. Marais, RUMM), www.rasch.org/andrich2009.htm

July 28-30, 2009, Tues.-Thur. PROMS HK 2009 Pacific Rim Objective Measurement Symposium, Hong Kong www.promshk.org

Sept. 6-11, 2009, Sun.-Fri. IMEKO XIX World Congress: Fundamental and Applied Metrology, Portugal www.imeko.org

Sept. 8, 2009, Tues. Rasch Refresher workshop (A. Tennant, RUMM), Leeds, UK, www.leeds.ac.uk/medicine/rehabmed/psychometric

April 30 - May 4, 2010, Fri.-Tues. AERA Annual Meeting, Denver, CO, USA, www.aera.net

June 14-15, 2010, Mon.-Wed. International Conference on Probabilistic Models for Measurement, Copenhagen, Denmark. www.rasch2010.cbs.dk

AERA 2009 Rasch-related Papers

American Educational Research Association Annual Meeting, San Diego, California, April 13-17, 2009

www.aera.net

Monday, April 13, 2009

Mon, Apr 13 - 12:00pm - 1:30pm - Omni San Diego, Gaslamp 1
Item Response Theory: Parametric and Nonparametric Models and Methods

An Extension of the MIRID Model for Polytomous Responses and Random Effects. Yongsang Lee, University of California - Berkeley; Mark R. Wilson, University of California - Berkeley

An Investigation of Cross-Classification Multilevel IRT Models. Paul T Ricci, University of Pittsburgh; Feifei Ye, University of Pittsburgh

Mon, Apr 13 - 2:15pm - 3:45pm - San Diego Convention Center, Room 5A
Methodology and Instrumentation for Mathematics Education Research

Using Mixture Rasch Models to Assess Middle Grades Teachers' Capacities to Reason About Arithmetic With Rational Numbers. Andrew G. Izsak, San Diego State University; Chandra H. Orrill, University of Georgia; Allan S. Cohen, University of Georgia; Rachael Eriksen Brown, University of Georgia

Examining Psychometric Properties of the Child Observation Record Using Rasch Analyses. Elena Malofeeva, High/Scope Educational Research Foundation; Zongping Xiang, High/Scope Educational Research Foundation; Marijata C. Daniel-Echols, High/Scope Educational Research Foundation

Mon, Apr 13 - 4:05pm - 4:45pm - San Diego Convention Center, Ballroom 6A
New Perspectives and Interventions in Literacy Instruction, Part 1

Fifth-Grade Reading Comprehension: An IRT Examination of Difficulty. Deni L Basaraba, University of Oregon; Julie Alonzo, University of Oregon; Gerald A. Tindal, University of Oregon

Mon, Apr 13 - 4:05pm - 5:35pm - San Diego Marriott Hotel & Marina, Marriott Hall Salon 3
Interesting Applications of Statistical and Psychometric Methods: Structured Poster Session

Efficient Common-Item Equating: Considering Spread of Item Difficulty. Anthony Daniel Albano, University of Minnesota

Tuesday, April 14, 2009

Tue, Apr 14 - 8:15am - 10:15am Building/Room: Omni San Diego / Gallery 3A
Rasch Issues of Dimensionality, Scaling, and Fit
Chair: Joy L. Matthews-Lopez, Prometric
Discussants: Leah Lyn Walker, University of California - Berkeley; Seock-Ho Kim, University of Georgia

A Comparison of Global Fit Indices as Indicators of Multidimensionality in Multidimensional Rasch Analysis. Leigh M. Harrell, Virginia Polytechnic Institute and State University; Edward W. Wolfe, Pearson

Effect of Ignoring Hierarchical Data Structures on Accuracy of Vertical Scaling: A Mixed-Effects Rasch Model Approach. Shudong Wang, ETS; Hong Jiao, University of Maryland; Ying Jin, American Institutes for Research

Measurement of Self-Authorship: A Validity Study Using Multidimensional Random Coefficients Multinomial Logit Model. Jessica Yue, Virginia Polytechnic Institute and State University; Elizabeth G. Creamer, Virginia Polytechnic Institute and State University; Edward W. Wolfe, Pearson

Strand-Based and Process-Based Multidimensionality and Rater Effects in Validation of the Carbon Cycle Learning Progression. Jinnie Choi, University of California - Berkeley; Yongsang Lee, University of California - Berkeley; Karen L. Draney, University of California - Berkeley

The Effects of Linking Procedure on the Scale Score. Brad Ching-Chao Wu, Pearson VUE; Huijuan Meng, Pearson VUE

Tue, Apr 14 - 10:35am - 11:15am - San Diego Marriott Hotel & Marina, Marriott Hall Salon 4
Division K: Teachers and Teacher Education, Session 18

An Exploratory Analysis of California Teaching Performance Assessment Task 3: An Application of the Rasch Measurement Model. Joseph Jesunathadas, California State University - San Bernardino

Tue, Apr 14 - 10:35am - 12:05pm - Omni San Diego, Gaslamp 2
New Dimensions in Early Literacy

A Rasch Analysis of Difficulty of Words in First-Grade Basal Readers for Preschool ELL and Monolingual English Speakers. Rebecca Deffes Silverman, University of Maryland - College Park; Cynthia B. Leung, University of South Florida; Ratna Nandakumar, University of Delaware

Tue, Apr 14 - 10:35am - 12:05pm - Omni San Diego, Gaslamp 3
Research in Adolescent and Adult Writing

Prominent Features in Writing Assessment and Implications for Instruction. Sherry Seale Swain, National Writing Project; Ricard L. Graves, Auburn University; David T. Morse, Mississippi State University

Tue, Apr 14 - 2:15pm - 3:45pm - Omni San Diego, Salon E
Equating and Scaling Issues

The Robustness of Rasch Preequating to Violations of Model Assumptions. Garron Gianopulos, University of South Florida

Tue, Apr 14 - 2:15pm - 3:45pm - San Diego Marriott Hotel & Marina, Irvine
Evaluation of Faculty Surveys and Students' Testing Experiences

Evaluating and Restructuring a New Faculty Survey: Measuring Perceptions Related to Research, Service, and Teaching. Kelly D. Bradley, University of Kentucky; Linda Worley, University of Kentucky; Jessica D. Cunningham, University of Kentucky; Jeffery P. Bieber, University of Kentucky

Tue, Apr 14 - 4:05pm - 5:35pm - Omni San Diego, Salon E
Computer Adaptive Testing Solutions

Using CAT to Achieve Comparability With a Paper Test. Tony D. Thompson, Pearson; Walter D. Way, Pearson

Tue, Apr 14 - 4:05pm - 6:05pm Building/Room: Omni San Diego / Balboa 1
Studies on Rasch Conditions and Applications
Chair: Nathaniel J. S. Brown, Indiana University - Bloomington
Discussants: Christian E. Mueller, The University of Memphis; G. Gage Kingsbury, Northwest Education Association

Analyzing Two-Tier Items With User-Defined Fit Statistics. Hak Ping Tam, National Taiwan Normal University; Margaret Li-min Wu, University of Melbourne

Calibration Methods Comparison With Rasch Model. Zhiming Yang, Pearson; Shudong Wang, ETS

Comparing Difficulty Estimates Using Operational CAT and Pilot Data. Timothy Joseph Muckle, AANA; James S. Masters, Pearson VUE / UNCG; Brian D. Bontempo, Mountain Measurement, Inc.

Evaluating Body of Work Judgments of Standard-Setting Panelists. Jade Caines, Emory University; George Engelhard, Emory University

On the Viability and Value of Invariant Reference Standards for Test- and Survey-Based Measures. William P. Fisher, Avatar International LLC

Tue, Apr 14 - 6:15pm - 7:45pm Building/Room: Omni San Diego / Gaslamp 1
Rasch Measurement SIG Business Meeting

Classical, IRT, and MIRT: A Primer on Test Equating. Mark H. Moulton, Educational Data Systems

Wednesday, April 15, 2009

Wed, Apr 15 - 9:05am - 9:45am - San Diego Convention Center, Ballroom 6A
Validation Studies of a Measure of Academic Resilience and the Big Five Personality Inventory

Fulfilling the Promise of Educational Research: Lessons From Plato and Aristotle. William P. Fisher, Avatar International LLC

Wed, Apr 15 - 10:35am - 11:15am - San Diego Convention Center, Ballroom 6A
Statistical Investigations of Multivariate Statistical Methods, Paper Discussion

Evaluation of Gibbs Sampling and Maximum Likelihood Methods Under the One-Parameter Logistic Model. Seock-Ho Kim, University of Georgia; Sukwoo Kim, Pusan National University; Daeyong Lee, Pusan National University

Wed, Apr 15 - 12:25pm - 1:55pm - Manchester Grand Hyatt, Madeleine Room B
Teachers' and Students' Physical Science Perspectives

> **Can Assessment of Student Conceptions of Force Be Enhanced Through Linguistic Simplification? A Rasch Model Common Person Equating of the FCI and the SFCI.** Sharon E. Osborn Popp, Arizona State University; Jane C. Jackson, Arizona State University

Wed, Apr 15 - 12:25pm - 1:55pm - San Diego Convention Center, Sails Pavilion
Advances in Motivation Research Poster Session

> **Understanding Self-Efficacy for Self-Regulated Learning in Mathematics: A Mixed Rasch Modeling Approach.** Ellen L. Usher, University of Kentucky; Jennifer Randall, University of Massachusetts

Wed, Apr 15 - 1:15pm - 1:55pm - San Diego Marriott Hotel & Marina, Marriott Hall Salon 4
Spirituality and Education: Roundtable Discussions

> **Mindfulness Practice in a University Classroom: A Kaleidoscope of Interdisciplinary Benefits.** Sharon G. Solloway, Bloomsburg University of Pennsylvania

Wed, Apr 15 - 2:15pm - 3:45pm - San Diego Marriott Hotel & Marina, Pacific
Collecting and Analyzing Survey Data

> **Constructing an Accurate Picture of Responses: A Comparison of Survey Results Through Measurement and Statistical Lenses.** Jessica D. Cunningham, University of Kentucky; Kelly D. Bradley, University of Kentucky

Wed, Apr 15 - 7:30pm - 9:30pm - San Diego Convention Center, Room 26B.
A joint-SIG social (Rasch SIG with other quantitative SIGs)

Thursday, April 16, 2009

Thu, Apr 16 - 8:15am - 10:15am Building/Room: Omni San Diego / Gallery 3A
Rasch Modeling for Cognitive Assessment
Chair: William P. Fisher, Avatar International LLC
Discussants: Brent M. Duckor, San Jose State University; Mary Garner, Kennesaw State University

> **Cognitive Assessment in Mathematics With the Least Squares Distance Method.** Lin Ma, University of Denver; Emre Cetin, Hacettepe University; Kathy E. Green, University of Denver
>
> **Conjunctive and Disjunctive Extensions of the Least Squares Distance Method, LSDM. for Cognitive Diagnosis.** Dimiter M. Dimitrov, George Mason University; Dimitar V. Atanasov, New Bulgarian University - Bulgaria
>
> **Coordinating Data Modeling Constructs**. Leah Lyn Walker, University of California - Berkeley; Robert Andrew Schwartz, University of California - Berkeley; David Torres Irribarra, University of California - Berkeley; Amy Dray, University of California; Mark R. Wilson, University of California - Berkeley
>
> **Performance of the Least Squares Distance Method, LSDM on Arithmetic Task Data With Observed Scores on Individual Attributes.** Sonia Janeth Romero, Universidad Autonoma de Madrid; Vicente Ponsoda, Universidad Autonoma; Dimiter M. Dimitrov, George Mason University; Xavier Giovanni Ordonez Camacho, Universidad Complutense de Madrid
>
> **Using Wright Maps to Understand Student Progress Toward College Readiness in Mathematics.** Diana Bernbaum Wilmot, University of California - Berkeley; Mark R. Wilson, University of California - Berkeley; Alan H. Schoenfeld, University of California; Danielle Champney, University of California - Berkeley; William Carl Zahner, University of California - Santa Cruz

Thu, Apr 16 - 10:35am - 11:15am - San Diego Marriott Hotel & Marina, Marriott Hall Salon 4
Meta-Analysis

> **Examining the Psychometric Properties of the Study Design and Implementation Device, DIAD.** Ryan Williams, Loyola University - Chicago; Therese D. Pigott, Loyola University - Chicago

Thu, Apr 16 - 10:35am - 12:05pm - Omni San Diego, Balboa 3
Factors Influencing Equating Accuracy

> **The Effects of Anchor Item Selection on IRT True Score Equating With Nonequivalent-Group Anchor-Test Design.** Hong Jiao, University of Maryland; Shudong Wang, ETS

Thu, Apr 16 - 11:25am - 12:05pm - San Diego Convention Center, Ballroom 6A
Measurement and Psychometric Topics, Paper Discussions

Assessing Unidimensionality in Item Response Data via Principal Component Analysis of Residuals From the Rasch Model. Mike McGill; Edward W. Wolfe, Pearson

Thu, Apr 16 - 12:25pm - 1:55pm - Omni San Diego, Salon E
Standards, Proficiency Judgments, and Norms

Establishing Criterion Measures on Graded Essays Using Objective Standard-Setting for Judge-Mediated Examinations. Gregory E. Stone, University of Toledo

Thu, Apr 16 - 3:05pm - 3:45pm Building/Room: San Diego Marriott Hotel & Marina / Marriott Hall Salon 4
Rasch Survey Analysis and Applications (Paper Discussions)

A Latent Trait Analysis of Higher Education Infrastructure in Russia. Anatoly Andreyevich Maslak, Slavyanskon-Kuban State Pedagogical Institute; Nikolaus Bezruczko, Measurement and Evaluation Consulting; Danilov Andrey, Slavyansk-on-Kuban Pedagogical Institute

Applying Rasch Model and Generalizability Theory to Study Modified Angoff Cut Scores for Reporting With Vertical Scales. Alvaro J. Arce-Ferrer, Pearson; Ze Wang, University of Missouri; Qing Xue, Harcourt Assessment Inc.

Checking Dimensionality in Rasch Measurement With Standardized Residual. Wen-Chung Wang, The Hong Kong Institute of Education; Yeh-Tai Chou, National Chung Cheng University

Development and Validation of the Survey of Internet Risk and Behavior. Robert K. Gable, Johnson & Wales University; Larry H. Ludlow, Boston College; Stacey L. Kite, Johnson & Wales University; D. Betsy McCoach, University of Connecticut

How Rasch Objective Measurement Can Make a Difference to the Results of Classical Statistical Analyses. Zongmin Kang, University of Toledo; Gregory E. Stone, University of Toledo; Svetlana A. Beltyukova, University of Toledo

Validation of Measures of the Quality of the Mentoring Experiences of New Teachers. Mike McGill, Edward W. Wolfe, Pearson; Denis W. Jarvinen, Strategic Measurement and Evaluation, Inc.

Utilizing the Rasch Model to Refine and Calibrate Items on the Tacit Knowledge Inventory for Principals, TKIP. Christian E. Mueller, The University of Memphis; Lisa D. Horton, Prairie View A&M University

"No Child Left Behind" Criticized Long in Advance

"... the tendency at present is very strong to provide means of measurement which are concerned somewhat closely with school achievements, and which can be used by teachers and others with little technical training. There is also a tendency, because of this need for a large number of measurements in the case of educational problems, to try to devise tests which can be scored by persons utterly devoid of judgment concerning the products in question.

"It would ill become the present writer to protest against these two tendencies; and they are intrinsically healthy. There is, however, a real danger in sacrificing soundness of principle and precision of result to the demand that we measure matters of importance and measure them without requiring elaborate technique or much time of the measurer. The danger is that the attention of investigators will be distracted from the problems of pure measurement for measurement's sake, which are a chief source of progress in measuring anything. Perhaps not even one person in a million need feel this passion, but for that one to cherish it and serve it is far more important than for him to devise a test which thousands of teachers will employ. Opposition, neglect, and misunderstanding will be much less disastrous to the work of quantitative science in education than a vast output of mediocre tests for measuring this, that and the other school product, of which a large percent are fundamentally unsound."

Edward L. Thorndike *(1918). The nature, purposes, and general methods of measurements of educational products. Chapter II in G.M. Whipple (Ed.), The 17th yearbook of the National Society for Study of Education. Part II. The Measurement of Educational Products. Bloomington, IL: Public School Publishing Co. p. 20-21.*
www.archive.org/details/measurementofedu00whiprich
Courtesy of Andrew Stephanou, ACER

Rasch Measurement Transactions
P.O. Box 811322, Chicago IL 60681-1322
www.rasch.org/rmt
Editor: John Michael Linacre

SIG Chair: Ed Wolfe, Secretary: Timothy Muckle
Program Chairs: Dimiter Dimitrov & Diana Bernbaum
SIG website: www.raschsig.org

BEN'S
"STABILITY THEOREM"

Rasch ability B_G of group G
to solve a problem of Rasch difficulty D is

STABLE

i.e. **ROBUST** against: interval merging, unbiased migration, multiple entry, joint effort and hierarchical regrouping

IF AND ONLY IF

$$B_G - D = \lambda(B_P - D) + (1 - \lambda)(B_T - D)$$

$$0 < \lambda < 1$$

$$(B_P - D) = \sum_n^N (B_n - D)/N + \log N$$

$$(B_T - D) = \sum_n^N (B_n - D)$$

Benjamin D. Wright's Handout, June 21, 1998. P=Pack, T=Team, N=Group size