

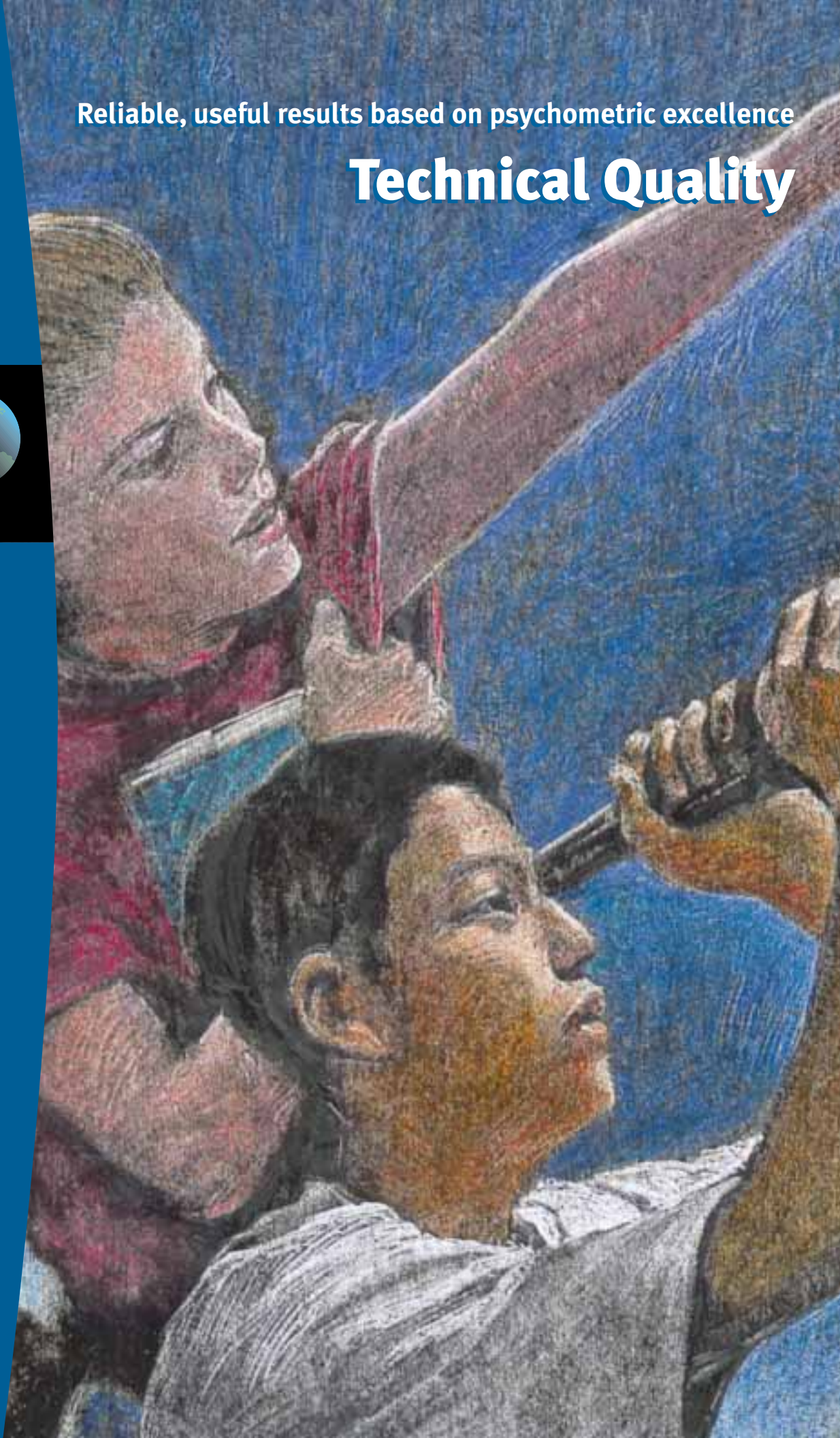
Reliable, useful results based on psychometric excellence

Technical Quality

TerraNova 

THE SECOND EDITION

 **CTB**
McGraw-Hill



EXCELLENCE IN ASSESSMENT:

Innovation combined with proven
research and development procedures.

Instructional **relevance**.

Advanced **measurement capabilities**.

This is

TerraNova, The Second Edition

—the newest member of the *TerraNova* family.

Important educational decisions require precise and accurate measurement information. The technical superiority of *TerraNova, The Second Edition* offers you significant advantages in today's challenging assessment environment.

Standard-Setting for *TerraNova, The Second Edition*



The carefully selected content and scientific scaling of *TerraNova, The Second Edition*, ensures that the assessment provides rich information about what students know and are able to do. Having this information summarized in a set of test scores, educators and parents want to know how student performance measures against the expectations set for them. This requires that we articulate these expectations as points on the achievement scale reported by the test. The process of articulating clearly defined expectations on a test's reporting scale is called standard-setting, and meaningful standards based reporting enriches the value of information provided by *TerraNova, The Second Edition*.

Setting standards on tests is a familiar activity for teachers. They articulate ranges of test scores (e.g., 90-100%, 80-89%, 70-79%, 60-69%, 0-59%) and associate these ranges with labels and descriptions (e.g., 'A, excellent;' 'B, good;' 'C, fair;', etc.). However, using these percentages to determine Performance Level placement regardless of the test being administered does not take into account the difficulty of the test or the specific knowledge, skills, and abilities required to answer the test questions.

Just as the content of *TerraNova, The Second Edition* reflects a scientific sampling of instructional content used across the nation, the standards set for *TerraNova, The Second Edition*, reflect a national consensus regarding standards determined by a broad range of educational experts. To achieve this consensus, CTB created a groundbreaking standard-setting method called the Bookmark Procedure (Lewis, Mitzel, & Green, 1996; Lewis, Green, Mitzel, Baum, & Patz, 1998).

The Bookmark Standard-Setting Procedure

In 1996, CTB research scientists developed a standard-setting method called the Bookmark Procedure. The fundamental tasks that participants perform under the Bookmark Procedure are:

- study the test items in order of difficulty;
- determine the increase in the knowledge, skills, and abilities required to respond to items as the items increase in difficulty;
- make judgments as to which items represent content that students in each Performance Level should know and be able to do; and
- write descriptions of expected student performance at each level.



In order to establish Performance Levels for *TerraNova* and *TerraNova, The Second Edition*, CTB relied upon the judgment of experts. Expert teachers nominated from across the nation participated in a Bookmark standard-setting to determine cut scores for five Performance Levels—Advanced, Proficient, Nearing Proficiency, Progressing, and Step 1/Starting Out. Teachers, parents and students can use information provided through the work of these experts to better understand the knowledge, skills, and abilities currently held by a student and which skills the student must yet attain in order to move into the next higher Performance Level.

The Bookmark Procedure has gained national prominence since its development in 1996. CTB conducted a professional training session at the 1999 annual meeting of the National Council for Measurement in Education in Montreal which was attended by university professors and staff from several test publishers who learned how to successfully implement CTB's Bookmark Procedure. Since its inception, the Bookmark Procedure has been used to set cut scores for at least 18 different states' assessment programs—13 of these standard-settings have been conducted by CTB.

The Bookmark Procedure is described in detail in various papers presented at national conferences, copies of which are available on request. CTB is very proud of this contribution to the field of measurement and the advantages the Bookmark Procedure offers to users of *TerraNova, The Second Edition*.





References

- Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks/Cole.
- Bock, E. Darrell (1997). A brief history of item response theory. *Educational Measurement: Issues and Practice*, 4, 21–33.
- Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 66, 443–459.
- Burket, G. R. (1988). ITEMSYS [Computer program]. Unpublished.
- Burket, G. R. (1991). PARDUX [Computer program]. Unpublished.
- Fitzpatrick, A. R., Link, V., Yen, W. M., Burket, G. R., Ito, K., & Sykes, R. (1996). Scaling performance assessments: A comparison of one-parameter and two-parameter partial credit models. *Journal of Educational Measurement*.
- Green, D. R., Yen, W. M., & Burket, G. R. (1989). Experiences in the application of item response theory in test construction. *Applied Measurement in Education*, 2, 297–312.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: John Wiley & Sons.
- Jolly, S. J., Johnson, R., Jones, B., & Abalus, J. (1986, April). *The effect of test speededness and random guessing on the validity of reading comprehension scores*. Paper presented at the annual meeting of the American Educational Research Association. San Francisco, CA.
- Lewis, D. M., Mitzel, H. C., Green, D. R. (1996). Standard Setting: A Bookmark Approach. In D. R. Green (Chair), IRT-Based Standard-Setting Procedures Utilizing Behavioral Anchoring. Symposium presented at the 1996 Council of Chief State School Officers 1996 National Conference on Large Scale Assessment, Phoenix, AZ.
- Lewis, D. M., Green, D. R., Mitzel, H. C., Baum, K., & Patz, R. J. (April, 1998). The Bookmark Standard Setting Procedure: Methodology and Recent Implementations. Paper presented at the 1998 National Council for Measurement in Education annual meeting, San Diego, CA.
- Lin, M. H. (1986, April). *The impact of time limits on test behaviors*. Paper presented at the annual meeting of the American Educational Research Association. San Francisco, CA.
- Linn, R. L., & Harnisch, D. (1981). Interactions between item content and group membership in achievement test items. *Journal of Educational Measurement*, 18, 109–118.
- Macmillan/McGraw-Hill. (1993). *Reflecting Diversity: Multicultural Guidelines for Educational Publishing Professionals*. New York, NY.
- McGraw-Hill. (1983). *Guidelines for Bias-free Publishing*. Monterey, CA.
- Munger, G. F., & Loyd, B. H. (1991). Effect of speededness on test performance of handicapped and nonhandicapped examinees. *Journal of Educational Research*, 85 (1), 53–57.
- Muraki, E. (1990). Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement*, 14, 59–71.
- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, 16, 159–176.
- Patz, R. J., and Junker, B. W. (1999). Applications and extensions of MCMC for IRT: Multiple item types, missing data, and rated responses. *Journal of Educational and Behavioral Statistics*, 24, 342–366.
- Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, 7, 201–210.
- Thissen, D. (1982). Marginal maximum likelihood estimation for the one-parameter logistic model. *Psychometrika*, 47, 175–186.
- Yen, W. M. (1993). Scaling performance assessments: Strategies for managing local item dependence. *Journal of Educational Measurement*, 30, 187–213.



The
Measure | **of**
Success

CTB/McGraw-Hill

A Division of The McGraw-Hill Companies



20 Ryan Ranch Road
Monterey, California 93940-5703
800/538-9547
www.ctb.com

CTB/McGraw-Hill Regional Offices

Western Regional Office

20 Ryan Ranch Road
Monterey, California 93940-5703
831/393-7780

Northern Regional Office

12412 Powerscourt Drive, Suite 110
St. Louis, Missouri 63131-9998
314/821-1700

Southern Regional Office

3260 Peachtree Industrial Boulevard, Suite 20
Duluth, Georgia 30096-2547
770/622-4300

Published by CTB/McGraw-Hill, a division of The McGraw-Hill Companies, Inc., 20 Ryan Ranch Road, Monterey, California 93940-5703. Copyright © 2001 by CTB/McGraw-Hill. All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

The materials shown in this publication are non-secure samples of secure test materials, some of which may be in unpublished form and under development by CTB. The final published version may have modifications.

California Achievement Tests, CTBS, SUPERA, and *TerraNova* are registered trademarks, and CAT and Classroom Connections are trademarks, of The McGraw-Hill Companies, Inc.