

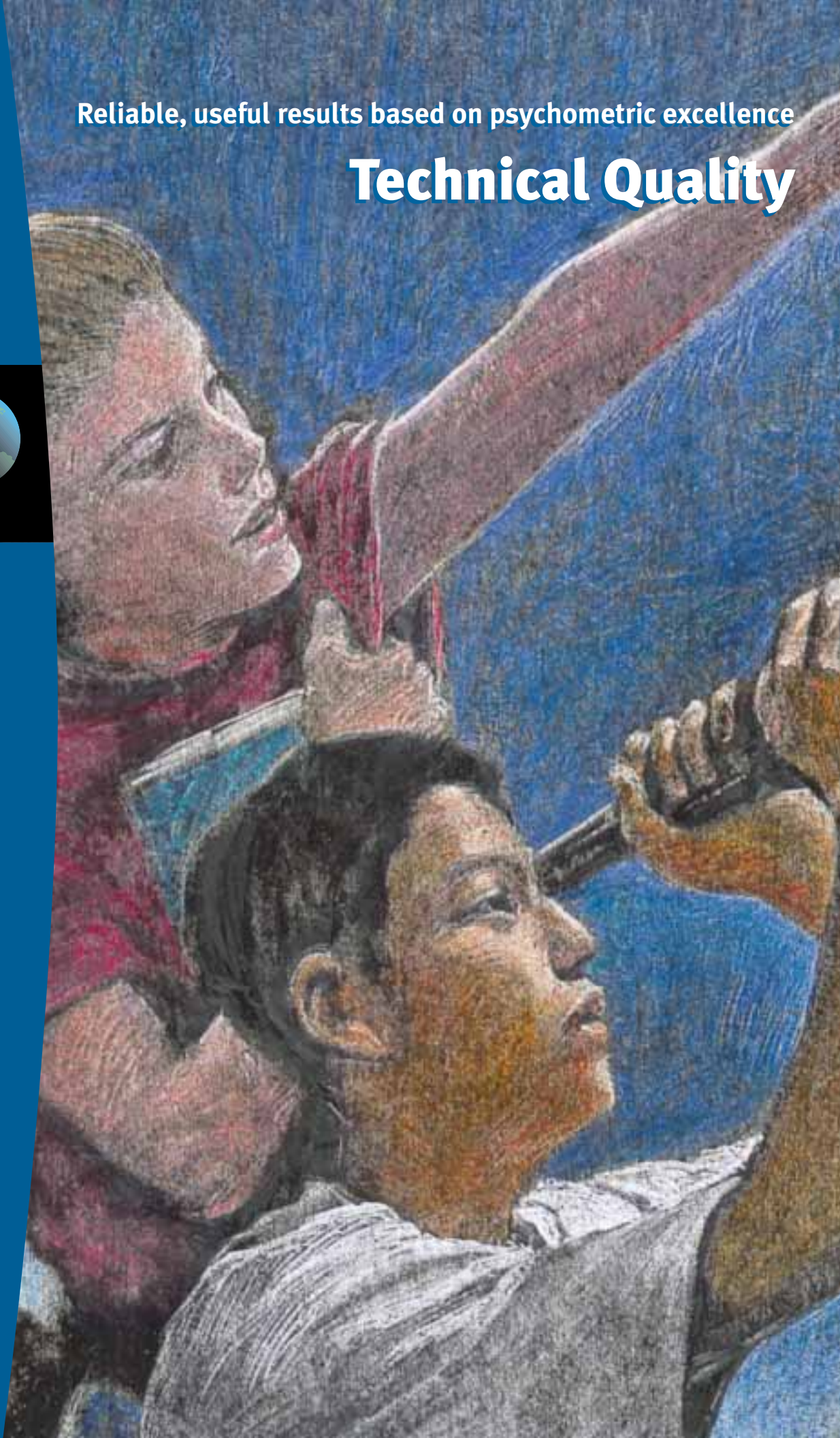
Reliable, useful results based on psychometric excellence

# 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.

# Scoring for *TerraNova, The Second Edition*



CTB's scoring services for *TerraNova, The Second Edition* reflect the same integrity and technical excellence that drove the development of this assessment series. Our regional scoring centers and highly trained staff ensure accurate, efficient processing of even the most complex *TerraNova* assessments. All sites are fully equipped with advanced hardware and software to ensure the highest possible standards of quality, service, and responsiveness to customer needs. We maintain stringent security and quality control for all test materials, specifications, and score reports. Because CTB manages the entire scoring and reporting process, each step is smoothly and efficiently integrated with the next.

## Advanced Scoring Systems

CTB leads the industry in advanced imaging—technology that captures electronic images of test book pages and routes them for computerized scoring. This electronic process provides highly accurate, consistent scoring for all assessments, even those including students' written work, such as essays and constructed-response math solutions. Evaluators electronically retrieve student responses and score them using on-screen rubrics. Image-based scoring provides the highest possible speed, consistency, and reliability for scoring multiple measures of student performance. Image-based results from constructed-response items can be electronically combined with selected-response test results to produce total scores for each assessment module.

CTB uses a range of advanced, real-time quality assurance techniques. We employ multiple readings of student responses to check consistency and guarantee accuracy in scores generated by different evaluators. Multiple, random read-behinds are conducted by CTB supervisors throughout the day to confirm that established scoring guidelines are being followed. Check-sets of previously-scored responses are also administered to verify correct application of scoring rules by evaluators and team leaders. This electronic management of test materials ensures that CTB's scoring processes deliver the most reliable, consistent results possible.



## Meaningful Scores

CTB reports three types of scores for *TerraNova, The Second Edition*:

- Norm-referenced scores,
- Criterion-referenced scores, and
- Performance Level scores.

### **Norm-Referenced Scores**

Norm-referenced scores describe individual student performance relative to the performance of a large, nationally representative group of students. CTB provides a wide variety of norm-referenced scores, including National Percentiles, Normal Curve Equivalents, Stanines, and Grade Equivalents.

### **Criterion-Referenced Scores**

Criterion-referenced scores are reported for each content objective measured by the test in terms of the Objectives Performance Index (OPI). The OPI is a unique score developed by CTB to provide more accurate objectives-level data for instructional planning and improvement. Because most objectives are measured by a relatively small number of items, CTB uses a Bayesian procedure that improves the reliability of the objective scores by taking the student's overall test performance into account. Essentially, the OPI is an estimate of the true score for an objective—the estimated proportion of total or maximum points possible—based on the overall test performance of a given student. The OPI is reported for each content objective measured.

The Technical Bulletin for *TerraNova, The Second Edition* provides a more complete explanation of the OPI. OPI scores are used to identify the level of mastery students have achieved for each objective. In particular, three levels of mastery are defined:

- Students are classified as having a high degree of mastery when their OPI for an objective is greater than or equal to the OPI expected for proficient students.
- Students who are classified as having a moderate degree of mastery demonstrate skills that approach, but do not reach, the performance expected of the proficient student.
- Students whose performance is below the moderate level are classified as having a low degree of mastery.

### **Performance Level Scores**

Performance Level scores provide a description of what students can do in terms of the content and skills assessed. These scores provide a means of comparing test results not with other students' results but with standards of academic performance that have been determined by panels of expert teachers, research scientists, and content specialists. Performance Level reporting establishes standards for students that are clearly delineated, that do not change over short periods of time, and that represent meaningful and significant learning. The Performance Levels are based on CTB's unique, widely respected Bookmark standard-setting procedure.



# 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.