

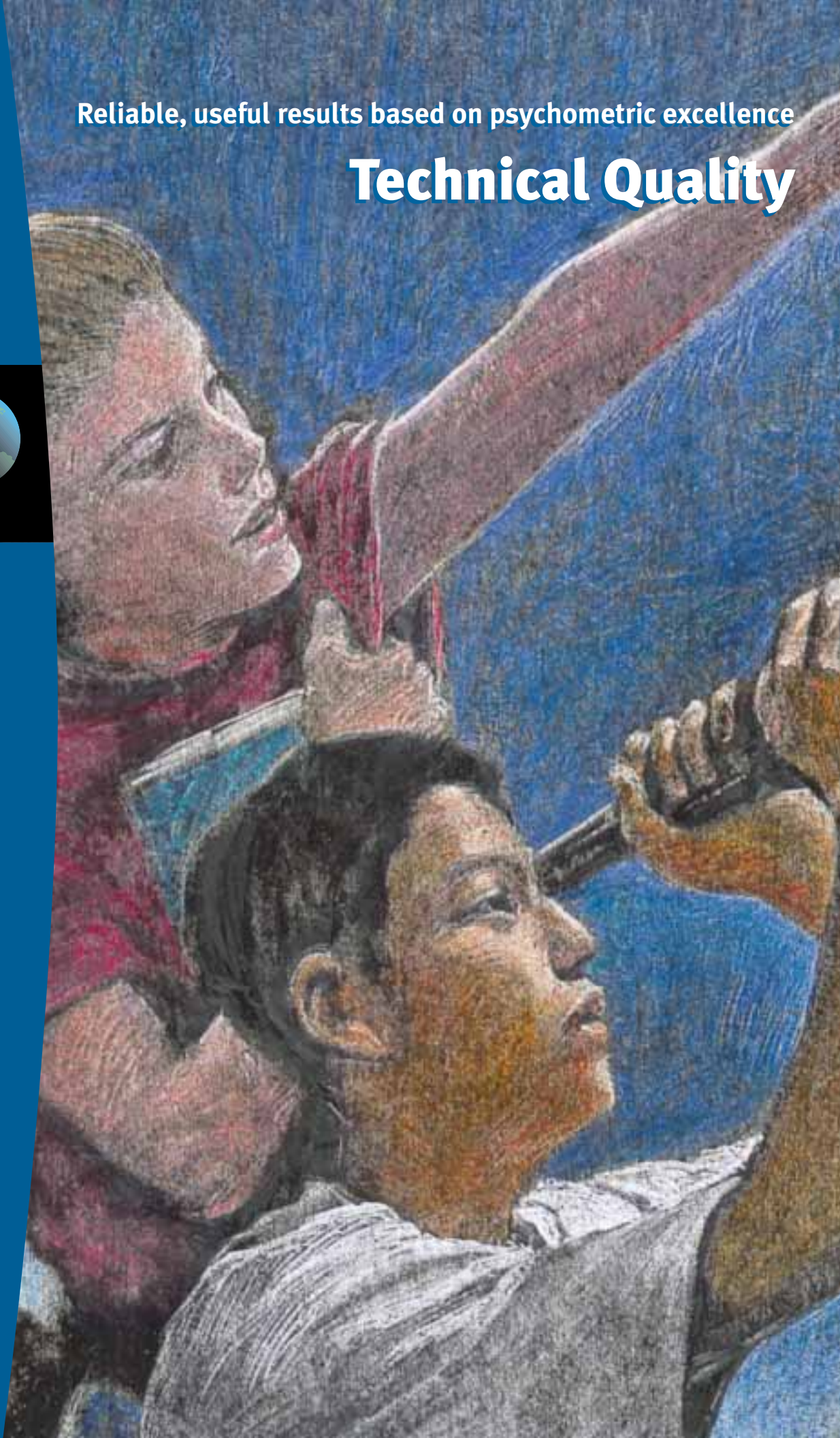
Reliable, useful results based on psychometric excellence

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.

Standardization of *TerraNova, The Second Edition*



National standardization for the final assessments was completed in October 1999 and April 2000. A winter study for Levels 10 and 11 was completed in February 2000.

Standardized or norm-referenced scores allow comparisons of students' performance with the performance of students in a specified group. National norms are developed so that students' scores can be compared with the performance of other students across the country.

In developing norms, it is not feasible to administer tests to all the students in the nation at all times of the year, so the tests are administered to a representative sample of students—the norming sample—at a particular time, the norming date. More than 275,000 students, from Kindergarten through Grade 12, participated in the standardization studies for this new edition of *TerraNova*.

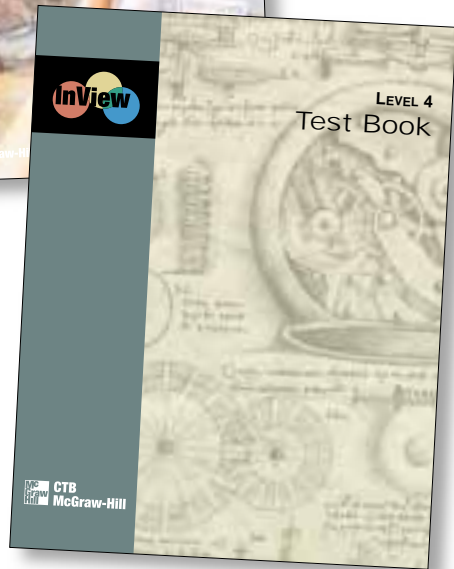
The 1999–2000 norms reflect a major redesign of CTB norming procedures. This redesign ensured collection of national norming information that was both more precise and more inclusive.

Greater precision has been achieved by sampling individual schools rather than districts, and by stratifying the nation's schools according to more detailed demographic information. For example, under the previous norming design we could assure that we selected an appropriate number of students from districts classified as urban, suburban, or rural. We are now able to select an appropriate number of students from schools located in the urban fringe of large central cities and at six other levels of urbanization.

Greater inclusiveness has been built into the sampling design to ensure that CTB's norms reflect the realities of today's testing programs. We have defined a standard school administration as one that involves all students according to their Individual Education Plans (IEPs), including

participation with testing accommodations as indicated. This definition—which complements the familiar definition of a standard student administration—has allowed CTB to collect meaningful, inclusive norming data in the standardization studies. More precise and inclusive norms will better meet the need for valid interpretations of the results of inclusive administrations of the new assessments.





Co-standardized Aptitude Assessment

CTB's newly developed test of cognitive skills, *InView*, was standardized concurrently with the new edition of *TerraNova* in the spring of 2000. The Primary Test of Cognitive Skills (PTCS) was also standardized concurrently with this edition. These tests measure students' thinking and reasoning skills by assessing performance on several cognitive tasks. PTCS, for Grade 1, measures verbal, spatial, memory, and conceptual abilities. *InView*, for Grades 2 through 12, measures verbal, quantitative, and other nonverbal reasoning skills that contribute to student success in educational programs.

When administered in combination with the second edition of *TerraNova*, *InView* and PTCS yield anticipated achievement scores that represent the most likely scores for students of similar attributes: chronological age in months, grade and month in school, and *InView* or PTCS scores. These scores may be viewed as a special kind of norm that enables users to compare an individual's level of achievement with students of similar age, grade, and cognitive abilities. The difference between a student's obtained and anticipated scores is an estimate of the student's achievement above or below the average of students with similar attributes, and can help screen students for a potential-actual discrepancy requiring further diagnostic testing.

Sampling Procedures

Students in the fall and spring norm groups were identified using stratified random sampling procedures. The design ensured that the norm group constitutes a sample of students that accurately represents the nation's school population, thereby fairly representing all minority and socioeconomic groups. To achieve the goal, CTB used two sampling designs. Public schools were stratified by region, community type, and socioeconomic status (Orshansky percentile). Catholic schools and private, non-Catholic schools were stratified by region and community type.

To ensure continuity in the norms from grade to grade, CTB used special sampling procedures for public schools and Catholic schools. Public and Catholic high schools were randomly sampled and asked to identify the major elementary and/or middle schools that send their students to the high school. We asked that students from these middle schools be included in the testing, and that elementary schools that contribute students to these middle schools also be included in the testing. We asked that all students be tested in the private, non-Catholic schools that participated in the standardization studies.

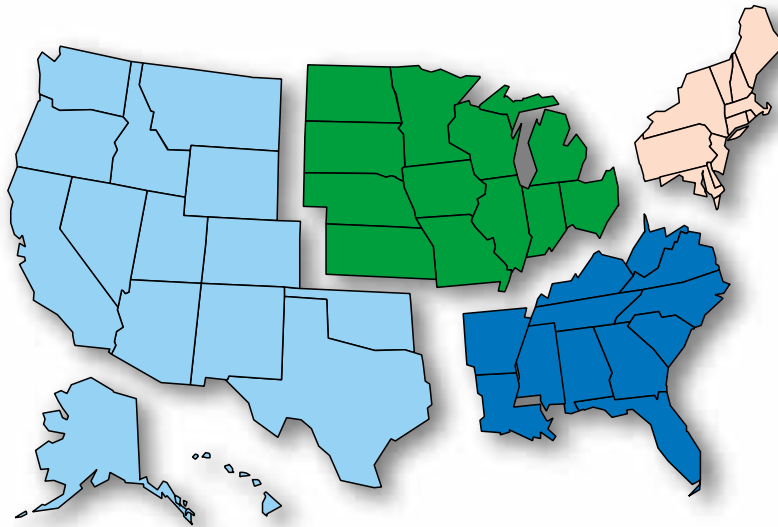
Stratification Variables

CTB drew on Quality Education Data, Inc. (QED), and their National Education Database™—a comprehensive census database of K–12 educational institutions in the United States—to obtain sampling data on geographic region, community type, socioeconomic status, and special needs. QED provides accurate and timely information about the nation's educational institutions. Use of these data helps to ensure that our norms accurately reflect achievement levels of the nation's student population as a whole.

Geographic Region

Four geographic regions constitute the first variable used for stratification during the standardization. These regions, and the states included in each region, are as follows:

- Eastern:
Connecticut, Delaware, District of Columbia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont
- Southern:
Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia
- Midcontinent:
Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin
- Western:
Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oklahoma, Oregon, Texas, Utah, Washington, Wyoming



Community Type

The second stratification variable—community type—is based on a seven-level metropolitan status classification system designated by the U.S. Bureau of the Census and recorded by QED. The levels identified included:

- Large central city
- Midsize central city
- Urban fringe of large central city
- Urban fringe of midsize central city
- Large town
- Small town
- Rural

Socioeconomic Status

Public schools are further stratified into two categories of socioeconomic status, high and low, based on the percentage of students eligible for Title I funding—that is, the percentage of children living in households with incomes below the federal poverty level. To assure further sample diversity, schools within cells were stratified by state and school size.

Special Needs

Standardization participants in the selected schools were asked to test all students who would ordinarily be included in their regular testing program, including those who would participate in regular testing with accommodations to meet special needs. To ensure an accurate description of the samples, each standardization examiner was asked to describe the category of special needs that any student might have, as well as any modified conditions for testing specified by the student's Individual Education Plan (IEP) that were applied during the standardization testing in each content area. This category was identified on the answer sheet for each student during each standardization. Special-needs categories for the standardization included:

- Learning-disabled
- Physically disabled
- Emotionally disabled
- Mentally disabled





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.