Technical Report 2402.2

An Authoritative Bibliography of Technical Adequacy Research Conducted on easyCBM – 2024 (Technical Report # 2402.2)

Denise Swanson
Gerald Tindal, PhD

University of Oregon

Published by

Behavioral Research and Teaching
University of Oregon • 175 Education
5262 University of Oregon • Eugene, OR 97403-5262
Phone: 541-346-3535 • Fax: 541-346-5689
http://brt.uoregon.edu

Note: This technical report was supported in part by Riverside Insight, the exclusive distributor of easyCBM©. The report does not reflect any endorsement by any of these organizations.

APA Reference: (Swanson & Tindal, 2024). Swanson, D., & Tindal, G. (2024). *An authoritative bibliography of technical adequacy research conducted on easyCBM – 2024 (Technical Report # 2402.2)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Abstract

This technical report provides an authoritative bibliographic resource of all the studies conducted on *easyCBM*® and published on the main website for Behavioral Research and Teaching under *Publications* (https://brtprojects.org). The *easyCBM*© software is a direct descendent of *Curriculum-based Measurement* (CBM) at the University of Minnesota, which has over 40 years of research supporting its technical adequacy. In this report, however, we only include research on the *easyCBM*© software itself, summarized in hundreds of technical reports (TRs), not on the larger body of research conducted on CBM more generally. Beyond this introduction and a brief discussion with three published summaries, this report is organized into three sections: (a) test development, (b) reliability, and (c) validity. Within each section, we provide a brief introduction highlighting the main types of research that have been completed on *easyCBM*© and then simply list all the technical reports (TR), available by placing the TR Number in the search key on the BRT *Publications* page. Each individual TR provides an abstract, brief introduction, and tables of results. Note that the TRs on test development and validation document a structured process not appropriate for the professional peer-reviewed literature because they are too detailed (with extensive tables for others to review) and do not contribute new insights and knowledge. Nevertheless, they need to be published and made publicly available. In obvious areas where test development and validation can be proffered with new insights for the professional literature, such publications appear in the form of peer-reviewed published articles, conference papers, and book chapters, attainable from the lead authors in BRT through their vita, also available on the main BRT website, https://brtprojects.org, Meet Us…About Our Team.

Introduction

We center our authority in educational testing on the standards published by American Educational Research Association (AERA) et al. (2014). As noted on the AERA website (https://www.apa.org/science/programs/testing/standards): They "represent the gold standard in guidance on testing in the United States and in many other countries." The standards, provide comprehensive not only on technical adequacy but also on highly related issues such as fairness in testing and training of professionals. In this overview, we first consider the TRs on test development and then provide a summary of the two most important topics—reliability and validity. Finally, we present three comprehensive references in a discussion, reflecting on this program of research and its support for decision-making.

Test Development

Though not explicitly addressed as a topic on its own in the *Standards*, such test development needs to be carefully considered in evaluating instruments for reliability and validity. In much of this research, we present on our use of item-response theory (IRT) to ensure the equivalence of forms so that benchmark and progress monitoring can be conducted in a seamless fashion. BRT was one of the first to use IRT in developing alternate forms, an improvement on the early days of CBM development that relied on using random sampling techniques to create alternate forms (an approach still used by DIBELS Next and DIBELS 8th Edition). This use of IRT is important because it allows the use of proper scales rather than raw scores (e.g., is the basis for development of an overall Reading Proficiency score, comprised of Proficient Reading, Vocabulary, and Passage Oral Reading Fluency). Note: In this bibliography, we also include the development of other measures, at times those used in the early development of easyCBM©s and at times those developed for validating easyCBM©s.

Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012a). *The development of CBM vocabulary measures: Grade 2 (Technical Report # 1209)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012b). *The development of CBM vocabulary measures: Grade 3 (Technical Report # 1210)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012c). *The development of CBM vocabulary measures: Grade 4 (Technical Report # 1211)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012d). *The development of CBM vocabulary measures: Grade 5 (Technical Report # 1212)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012). *The development of CBM vocabulary measures: Grade 6 (Technical Report # 1213)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012e). *The development of CBM vocabulary measures: Grade 7 (Technical Report # 1214)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Anderson, D., Park, B. J., & Tindal, G. (2012f). *The development of CBM vocabulary measures: Grade 8 (Technical Report # 1215)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Gonzalez, M., & Tindal, G. (2013). *The development of easyCBM Spanish literacy assessments for use in Grades K-2 (Technical Report # 1301)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Lai, C. F., Anderson, D., Park, B. J., & Tindal, G. (2012). *An examination of test-retest, alternate form reliability, and generalizability theory study of the easyCBM reading assessments: Grade 4 (Technical Report # 1219)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Lai, C. F., & Tindal, G. (2009a). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 3 (Technical Report # 0902)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Lai, C. F., & Tindal, G. (2009b). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 4 (Technical Report # 0903)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Lai, C. F., & Tindal, G. (2009c). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and the general education populations: Grade 2 (Technical Report # 0920)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Park, B. J., & Tindal, G. (2008). *The development of middle school passage reading fluency measures in a progress monitoring assessment system (Technical Report # 46)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Park, B. J., & Tindal, G. (2012a). *The development of the easyCBM CCSS reading assessments: Grade 3 (Technical Report # 1221)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Park, B. J., & Tindal, G. (2012b). *The development of the easyCBM CCSS reading assessments: Grade 4 (Technical Report # 1222)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Park, B. J., & Tindal, G. (2012c). *The development of the easyCBM CCSS reading assessments: Grade 5 (Technical Report # 1223)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Park, B. J., & Tindal, G. (2012d). *The development of the easyCBM CCSS reading assessments: Grade 6 (Technical Report # 1224)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Park, B. J., & Tindal, G. (2012). *The development of the easyCBM CCSS reading assessments: Grade 7 (Technical Report # 1225)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Park, B. J., & Tindal, G. (2012e). *The development of the easyCBM CCSS reading assessments: Grade 8 (Technical Report # 1226)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2007a). *The development of early literacy measures for use in a progress monitoring assessment system: Letter names, letter sounds and phoneme*

segmenting (Technical Report # 39). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2007b). *The development of word and passage reading fluency measures in a progress monitoring assessment system (Technical Report # 40)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2008). *The development of fifth-grade passage reading fluency measures for use in a progress monitoring assessment system (Technical Report # 43)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2009a). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 1 (Technical Report # 0919)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2009b). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Kindergarten (Technical Report # 0921)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Tindal, G., Ulmer, K., & Glasgow, A. (2006). *easyCBM® online progress monitoring assessment system* http://easycbm.com. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Alonzo, J., Tindal, G., Farley, D., Irvin, P. S., Lai, C. F., Saven, J. L., & Wray, K. A. (2014). *Technical manual: easyCBM (Technical Report # 1408)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Irvin, P. S., Patarapichayatham, C., Alonzo, J., & Tindal, G. (2012). *The development and scaling of the easyCBM CCSS middle school mathematics measures (Technical Report # 1207)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Irvin, P. S. (2014). *Measurement development and item alignment*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Irvin, P. S., Saven, J. L., Alonzo, J., Park, B. J., Anderson, D., & Tindal, G. (2013). *The development and scaling of the easyCBM CCSS elementary mathematics measures: Grade K (Technical Report # 1314)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Irvin, P. S., Saven, J. L., Alonzo, J., Park, B. J., & Tindal, G. (2013a). *The development and scaling of the easyCBM CCSS elementary mathematics measures: Grade 2 (Technical Report # 1316)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Irvin, P. S., Saven, J. L., Alonzo, J., Park, B. J., & Tindal, G. (2013b). *The development and scaling of the easyCBM CCSS elementary mathematics measures: Grade 4 (Technical Report # 1318)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Jung, E., Liu, K., Ketterlin-Geller, L. R., & Tindal, G. (2008). *Instrument development procedures for mathematics measures (Technical Report # 0802)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Alonzo, J., & Tindal, G. (2009a). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade 5 (Technical Report # 0901)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Alonzo, J., & Tindal, G. (2009b). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and general education populations: Grade*

8 (Technical Report # 0904). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Alonzo, J., & Tindal, G. (2009c). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and general populations: Grade 7 (Technical Report # 0908)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Alonzo, J., & Tindal, G. (2009d). *The development of K-8 progress monitoring measures in mathematics for use with the 2% and the general education populations: Grade 6 (Technical Report # 0907)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Liu, K., Ketterlin-Geller, L. R., Yovanoff, P., & Tindal, G. (2008). *Examining item functioning of math screening measures for grades 1-8 students (Technical Report # 0804)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Liu, K., Sundstrom-Hebert, K., Ketterlin-Geller, L. R., & Tindal, G. (2008a). *Instrument development procedures for maze measures (Technical Report # 0806)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Liu, K., Sundstrom-Hebert, K., Ketterlin-Geller, L. R., & Tindal, G. (2008b). *Instrument development procedures for silent reading measures (Technical Report # 0803)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Nese, J. F. T., Anderson, D., Hoelscher, K., Tindal, G., & Alonzo, J. (2011). *Progress mmonitoring instrument development: Silent reading fluency, vocabulary, and reading comprehension (Technical Report # 1110).* Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Park, B. J., Alonzo, J., & Tindal, G. (2011). *The development and technical adequacy of seventh-grade reading comprehension measures in a progress monitoring assessment (Technical Report # 1102)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Saven, J. L., Irvin, P. S., Park, B. J., Alonzo, J., & Tindal, G. (2013a). *The development and scaling of the easyCBM CCSS elementary mathematics measures: Grade 3 (Technical Report # 1317)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Saven, J. L., Irvin, P. S., Park, B. J., Alonzo, J., & Tindal, G. (2013b). *The development and scaling of the easyCBM CCSS elementary mathematics measures: Grade 5 (Technical Report # 1319)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Saven, J. L., Irvin, P. S., Park, B. J., Tindal, G., & Alonzo, J. (2013). *The development and scaling of the easyCBM CCSS elementary mathematics measures: Grade 1 (Technical Report # 1315)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Reliability

Reliability in educational testing refers to the consistency and stability of test scores over time, and across different forms, as well as within the tests themselves. Five types of reliability are considered to ensure that a test consistently measures what it is intended to, necessary to provide confidence in outcomes (scores and decisions).

Test-Retest Reliability. The emphasis in this type of reliability is the consistency of scores over time. To analyze test-retest reliability, the same test is administered to the same group of individuals on two different occasions, and the scores are then correlated. High test-retest

reliability indicates that the test produces stable results over time. This type of reliability is particularly critical when assessing dimensions that are expected to remain stable, such as intelligence or personality traits. Note: The interval of time is important and when applied to progress monitoring systems that are part of Response to Intervention (RTI), test-retest stability needs to be time-limited (e.g., one week) because the purpose of frequent measurement is to show growth, which would then be confounded as unreliability.

Inter-Rater Reliability. This form of reliability examines agreement between different raters or observers assessing the same response and is important in 'production' items (versus 'selection items' such as multiple-choice), where different raters must judge the outcome on a dimension of quality. High inter-rater reliability indicates that the scoring criteria are clear and consistent across different raters, reducing the risk of subjective bias.

Intra-Rater Reliability. This form of reliability is like inter-rater but assesses the consistency of a single rater's assessments across multiple occasions. Intra-rater reliability is important in situations where the same individual repeatedly evaluates multiple production responses. High intra-rater reliability suggests that the rater is consistent in their evaluations, minimizing variability due to subjective factors.

Parallel (Alternate)-Forms Reliability. This type of reliability assesses the consistency of scores between two equivalent versions of a test. To analyze alternate-forms reliability, both forms are administered to the same individuals, and their scores are correlated. High parallel-forms reliability indicates that both versions measure the same construct equally well. This method is particularly important in using a pre-test post-test design to demonstrate effects and mitigate against memory or practice.

Internal Consistency Reliability. Several reliability indices can be considered here, all of them focusing on the extent to which items within a test are consistent in measuring the same construct.

1. Cronbach's Alpha is the most common measure of internal consistency. It calculates the average correlation of all possible split-halves of the test. Values range from 0 to 1, with higher values indicating greater internal consistency.
2. Split-Half Reliability is a cruder form of Cronbach's Alpha: The test is divided into two halves (odd-even items or first-second half), and the scores for each 'half' are correlated. This method assumes that both halves are equivalent in measuring the construct and high correlation indicates good internal consistency.
3. Kuder-Richardson Formula 20 (KR-20) provides an estimate of the test's internal consistency based on the proportion of correct and incorrect answers and is like Cronbach's Alpha. KR-20 is used for dichotomous items (e.g., true/false questions).

In summary, reliability encompasses various methods to ensure that test scores are consistent over time, across different forms, and within the test itself. Test-retest reliability, inter-rater reliability, intra-rater reliability, parallel-forms reliability, and internal consistency each provide different types of evidence on the stability and consistency of test scores, contributing to the overall reliability which is necessary for validity in using educational assessments for making decisions. Note that reliability is necessary but not sufficient in educational and psychological

testing because precision and trustworthiness of test scores do not tell the whole story. Unreliable tests are likely to result in erroneous conclusions and decisions, affecting educational placements, psychological diagnoses, and research outcomes. By documenting various types of reliability, test developers and researchers can identify and address potential sources of error, enhancing the overall quality and effectiveness of their assessments. Following are the studies on easyCBM© that address reliability. In these references, the full range of reliability types is represented. Note: Some references are listed in both reliability and validity, as the technical report includes information on both (e.g., those with '*the technical adequacy' in the title).*

Alonzo, J., Anderson, D., & Tindal, G. (2009). *IRT analysis of general outcome measures in Grades 1-8 (Technical Report # 0916)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Liu, K., & Tindal, G. (2007). *Examining the technical adequacy of reading comprehension measures in a progress monitoring assessment system (Technical Report # 41)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Liu, K., & Tindal, G. (2008). *Examining the technical adequacy of second-grade reading comprehension measures in a progress monitoring assessment system (Technical Report # 0808)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2008). *Examining the technical adequacy of fifth-grade reading comprehension measures in a progress monitoring assessment system (Technical Report # 0807)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2009). *Alternate form and test-retest reliability of easyCBM reading measures (Technical Report # 0906)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Alonzo, J., & Tindal, G. (2013). *Study of the reliability of CCSS-aligned math measures (2012 research version): Grades 6-8 (Technical Report # 1312)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Lai, C. F., Park, B. J., Alonzo, J., & Tindal, G. (2012). *An examination of test-retest, alternate form reliability, and generalizability theory study of the easyCBM reading assessments: Grade 2 (Technical Report # 1217)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Park, B. J., Lai, C., F., Alonzo, J., & Tindal, G. (2012). *An examination of test-retest, alternate form reliability, and generalizability theory study of the easyCBM reading assessments: Grade 1 (Technical Report # 1216)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Tindal, G., & Alonzo, J. (2009). *Internal consistency of general outcome measures in grades 1-8 (Technical Report # 0915)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Guerreiro, M., Alonzo, J., & Tindal, G. (2014). *Internal consistency of the easyCBM CCSS reading measures grades K-8 (Technical Report # 1407)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Irvin, P. S., Alonzo, J., Lai, C. F., Park, B. J., & Tindal, G. (2012). *Analyzing the reliability of the easyCBM reading comprehension measures: Grade 7 (Technical Report # 1206)*. Eugegne, OR: Behavioral Research and Teaching, University of Oregon.

Irvin, P. S., Alonzo, J., Park, B. J., Lai, C. F., & Tindal, G. (2012). *Analyzing the reliability of the easyCBM reading comprehension measures: Grade 6 (Technical Report # 1205)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Irvin, P. S., Alonzo, J., Park, B. J., & Tindal, G. (2012). *Analyzing the reliability of the easyCBM reading comprehension measures: Grade 2 (Technical Report # 1201)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Irvin, P. S., Park, B. J., Alonzo, J., & Tindal, G. (2012). *Analyzing the reliability of the easyCBM reading comprehension measures: Grade 3 (Technical Report # 1202)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Park, B. J., Anderson, D., Alonzo, J., & Tindal, G. (2012). *An examination of test-retest, alternate form reliability, and generalizability theory study of the easyCBM reading assessments: Grade 5 (Technical Report # 1220)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Nese, J. F. T., Anderson, D., Irvin, P. S., & Alonzo, J. (2018a). *In-Brief: Reliability of the slope of the easyCBM® math measures (Technical Report # 1804)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Nese, J. F. T., Anderson, D., Irvin, P. S., & Alonzo, J. (2018b). *In-Brief: Reliability of the slope of the easyCBM® reading measures (Technical Report # 1803)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Park, B. J., Anderson, D., Alonzo, J., Lai, C. F., & Tindal, G. (2012). *An examination of test-retest, alternate form reliability, and generalizability theory study of the easyCBM reading assessments: Grade 3 (Technical Report # 1218)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Park, B. J., Irvin, P. S., Alonzo, J., Lai, C. F., & Tindal, G. (2012). *Analyzing the reliability of the easyCBM reading comprehension measures: Grade 4 (Technical Report # 1203)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Park, B. J., Irvin, P. S., Lai, C. F., Alonzo, J., & Tindal, G. (2012). *Analyzing the reliability of the easyCBM reading comprehension measures: Grade 5 (Technical Report # 1204)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Patarapichayatham, C., Anderson, D., Irvin, P. S., Kamata, A., Alonzo, J., & Tindal, G. (2011). *easyCBM slope reliability: Letter names, word reading fluency, and passage reading fluency (Technical Report # 1111)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Tindal, G. (2017). *Oral reading fluency: Outcomes from 30 years of research (Technical Report # 1701)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Wray, K. (2014). *Reliability evidence*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Wray, K., Alonzo, J., & Tindal, G. (2014a). *Internal consistency of the easyCBM CCSS math measures Grades K-8 (Technical Report # 1405)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Wray, K., Alonzo, J., & Tindal, G. (2014b). *Internal consistency of the easyCBM vocabulary measures Grades 2-8 (Technical Report # 1406)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Wray, K., Lai, C. F., Sáez. L., Alonzo, J., & Tindal, G. (2014). *easyCBM beginning reading measures: Grades K-1 alternate form reliability and criterion validity with the SAT-10*

(Technical Report # 1403). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Validity

In educational testing, validity refers to the degree to which evidence and theory support the interpretations of test scores for an intended purpose. Four types of validity ensure that a test measures what it is intended to measure and that the inferences made from test scores are appropriate: Each type of validity depends upon the type of decision being made. Note: We do not address face validity, which is not a formal part of the validation scheme. It simply refers to the extent to which a test appears to measure what it is supposed to measure, based on a superficial analysis (often formed by ill-defined expectations). Although face validity is not a rigorous form of validity, it may be important for gaining acceptance from test-takers and stakeholders.

Construct Validity. This type of validity is critical when referring to abstract 'dimensions' like 'achievement', 'personality traits', and 'intelligence'. Construct validity evaluates how well a test measures the theoretical construct it is intended to measure. This type of validity involves gathering evidence from multiple sources, including correlations with other measures, factor analysis, and outcomes from experimental interventions.

Content Validity. In this type of validity, the focus is on whether a test comprehensively covers the domain it is intended to measure. Typically, experts in the field assess content by reviewing items and ensuring they reflect the intended content area. Importantly, both breadth and depth need to be considered to ensure the items represent the entire range of possible content in the domain. Two types of invalidity may result: Construct under-representation or construct over-representation, which are interpreted as their names imply.

Criterion-Related Validity. Several types of criterion-related validity are possible, with two types of relations between the central measure and other related measures: (a) the time frame (whether prior, which would be *predictive*, or at the same time, which would be *concurrent*), and (b) whether the relation should be high (*convergent*) or low (*discriminate*). For example, fall benchmark CBMs can be used to predict end-of year performance on a state test (*predictive*), or the spring benchmarks can be used to determine if the CBM ranks the students in a similar manner as the state test administered in the same timeframe (*concurrent*). And, in either of these relations between CBMs and state tests, it would be expected that reading CBMs would be highly related to English language arts state test (*convergent*) but not math state test (*discriminate*). In turn, the same would be true for CBM math measures, which would be more highly related to the state math test than the English language arts test.

Consequential Validity. This form of validity is more current and involves evaluating the intended and unintended consequences of test use. Consequential validity considers the broader impact of test scores on individuals and society. It is established using a wide range of data from studies that may employ descriptive, quasi-experimental, or experimental designs. Furthermore, various outcomes need to be documented from samples of respondents (students, teachers, administrators, etc.), both in their personal demographics as well as in their responses to different measures and these data need to be analyzed using a range of statistical techniques.

In summary, validity ensures that a test measures what it is intended to measure and supports appropriate interpretations of test scores. Construct validity, content validity, criterion-related validity (predictive-concurrent and convergent-discriminant), and consequential validity each provide unique insights into the appropriateness of test interpretations. Establishing validity involves a combination of expert judgment, statistical analyses, and empirical evidence. The most important consideration of validity is that it is multi-dimensional and involves theory, evidence from a variety of sources (experts, other measures), and use of the measures in decision-making. The following references present the full range of validity studies conducted to date on *easyCBM*®. As with reliability, these studies deploy a wide range of evidence types to support use of *easyCBM*® in making a variety of decisions, from screening for risk to diagnostic analysis of performance and progress monitoring.

Alonzo, J. (2019). *Initial norms for the easyCBM® Spanish language vocabulary assessments (Technical Report # 1901)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Anderson, D. (2018a). *Classification accuracy of the easyCBM Kindergarten – Grade 2 reading measures (Technical Report # 1801)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Anderson, D. (2018b). *Supplementary report on easyCBM Grade K – 2 literacy measures: Classification accuracy and predictive validity, a follow-up to Technical Report # 1403 (Technical Report # 1805)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Anderson, D. (2018c). *Supplementary report on easyCBM MCRC Measures: A Follow-up to previous technical report (Technical Report # 1807)*. Eugene OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Anderson, D. (2018d). *Supplementary report on easyCBM PRF measures: A Follow-up to previous technical reports (Technical Report #1806)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Liu, K., & Tindal, G. (2007). *Examining the technical adequacy of reading comprehension measures in a progress monitoring assessment system (Technical Report # 41)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Liu, K., & Tindal, G. (2008). *Examining the technical adequacy of second-grade reading comprehension measures in a progress monitoring assessment system (Technical Report # 0808)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Park, B. J., & Tindal, G. (2013a). *An examination of the internal structures of the easyCBM CCSS reading measures (Technical Report # 1304)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Park, B. J., & Tindal, G. (2013b). *An examination of the internal structures of the Grade K-5 easyCBM CCSS reading measures: A Construct validity study (Technical Report # 1305)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Park, B. J., & Tindal, G. (2013c). *Examining the internal structure of the easyCBM reading measures, Grades K-5 (Technical Report # 1302)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2004a). *Analysis of reading fluency and comprehension measures for first-grade students (Technical Report # 25)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2004b). *Analysis of reading fluency and comprehension measures for fourth-grade students (Technical Report # 27)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2004c). *Analysis of reading fluency and comprehension measures for seventh-grade students (Technical Report # 23)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2004d). *Analysis of reading fluency and comprehension measures for sixth-grade students (Technical Report # 24)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2004e). *District reading assessments, spring 2004 administration (Technical Report # 30)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2008). *Examining the technical adequacy of fifth-grade reading comprehension measures in a progress monitoring assessment system (Technical Report # 0807)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., Tindal, G., & Lai, C. F. (2010). *Making the good even better: Feedback from easyCBM focus groups, school year 2009/2010 (Technical Report # 1001)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D. (2013). *Hierarchical linear modeling (HLM): An Introduction to key concepts within cross-sectional and growth modeling frameworks (Technical Report # 1308)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Alonzo, J., & Tindal, G. (2010a). *Diagnostic efficiency of easyCBM math: Oregon (Technical Report # 1009)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Alonzo, J., & Tindal, G. (2010b). *Diagnostic efficiency of easyCBM mathematics: Washington state (Technical Report # 1008)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Alonzo, J., & Tindal, G. (2010c). *easyCBM mathematics criterion related validity evidence: Oregon state test (Technical Report # 1011)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Alonzo, J., & Tindal, G. (2010d). *easyCBM mathematics criterion related validity evidence: Washington state test (Technical Report # 1010)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Alonzo, J., & Tindal, G. (2011a). *A cross-validation of easyCBM mathematics cut scores in Oregon: 2009-2010 (Technical Report # 1104)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Alonzo, J., & Tindal, G. (2011b). *A cross-validation of easyCBM mathematics cut scores in Washington state: 2009-2010 Test (Technical Teport # 1105)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Alonzo, J., & Tindal, G. (2011c). *easyCBM reading criterion related to validity evidence: Washington state test 2009-2010 (Technical Report # 1101)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Alonzo, J., & Tindal, G. (2011d). *easyCBM reading criterion related validity evidence: Oregon state test 2009-2010 (technical Report # 1103)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Alonzo, J., & Tindal, G. (2013). *easyCBM CCSS math item scaling and test form revision (2012-2013): Grades 6-8 (Technical Report # 1313)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Irvin, P. S., Alonzo, J., & Tindal, G. (2012). *The alignment of the easyCBM middle school mathematics CCSS measures to the Common Core State Standards (Technical Report # 1208)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Lai, C. F., Nese, J. F. T., Park, B. J., Sáez, L., Jamgochian, E. M., Alonzo, J., & Tindal, G. (2010). *Technical adequacy of the easyCBM primary-level mathematics measures (Grades K-2), 2009-2010 version (Technical Report # 1006)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Park, B. J., Irvin, P. S., Alonzo, J., & Tindal, G. (2011). *Diagnostic efficiency of easyCBM reading: Washington state (Technical Report # 1107)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Park, S., Alonzo, J., & Tindal, G. (2015). *An exploration of differential item functioning with the easyCBM middle school mathematics tests: Grades 6-8 (Technical Report # 1501)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Rowley, B., Alonzo, J., & Tindal, G. (2014). *Criterion validity evidence for the easyCBM CCSS math measures: Grades 6-8 (Technical Report # 1402)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Baker, D. L., Biancarosa, G., Park, B. J., Bousselot, T., Smith, J.-L., Baker, S. K., Kame' enui, E. J., Alonzo, J., & Tindal, G. (2015). Validity of CBM measures of oral reading fluency and reading comprehension on high-stakes reading assessments in grades 7 and 8. *Reading and Writing: An Interdisciplinary Journal*, *28*(1), 57-104. https://doi.org/DOI 10.1007/s11145-014-9505-4

Beghetto, R., & Tindal, G. (2004). *Analysis of reading fluency and comprehension measures for fifth grade students (Technical Report # 28)*. Behavioral Research and Teaching, University of Oregon.

Irvin, P. S., Park, B. J., Alonzo, J., & Tindal, G. (2012a). *The alignment of the easyCBM Grades 6-8 math measures to the Common Core Standards (Technical Report # 1230)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Irvin, P. S., Park, B. J., Alonzo, J., & Tindal, G. (2012b). *The alignment of the easyCBM Grades K-2 math measures to the Common Core Standards (Technical Report # 1228)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Irvin, P. S., Park, B. J., Anderson, D., Alonzo, J., & Tindal, G. (2011). *A cross-validation of easyCBM reading cut scores in Washington: 2009-2010 (Technical Report # 1109)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Jamgochian, E. M., Park, B. J., Nese, J. F. T., Lai, C. F., Sáez, L., Anderson, D., Alonzo, J., & Tindal, G. (2010). *Technical adequacy of the easyCBM Grade 2 reading measures (Technical Report # 1004)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Ketterlin-Geller, L. R., & Tindal, G. (2004). *Analysis of reading fluency and comprehension measures for 3rd-grade students (Technical Report # 22)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Alonzo, J., & Tindal, G. (2013a). *easyCBM reading criterion related validity evidence: Grades 2-5 (Technical Report # 1310)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Alonzo, J., & Tindal, G. (2013b). *easyCBM reading criterion related validity evidence: Grades K-1 (Technical Report # 1309)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Lai, C. F., Nese, J. F. T., Jamgochian, E. M., Kamata, A., Anderson, D., Park, B. J., Alonzo, J., & Tindal, G. (2010). *Technical adequacy of the easyCBM primary-level reading measures (Grades K-1), 2009-2010 Version. (Technical Report # 1003)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon..

Martinez, M., Ketterlin-Geller, L. R., & Tindal, G. (2007). *Content-related evidence for validity for mathematics tests: Teacher review (Technical Report # 42)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Nese, J. F. T., Lai, C. F., Anderson, D., Jamgochian, E. M., Kamata, A., Saez, L., Park, B. J., Alonzo, J., & Tindal, G. (2010). *Technical adequacy of the easyCBM mathematics measures: Grades 3-8: 2009-2010 version (Technical Report # 1007)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Nese, J. F. T., Lai, C. F., Anderson, D., Park, B. J., Tindal, G., & Alonzo, J. (2010). *The alignment of easyCBM math measures to curriculum standards (Technical Report #1002)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Park, B. J., Anderson, D., Irvin, P. S., Alonzo, J., & Tindal, G. (2011). *Diagnostic efficiency of easyCBM reading: Oregon (Technical Report # 1106)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Park, B. J., Irvin, P. S., Alonzo, J., & Tindal, G. (2012). *The alignment of the easyCBM grades 3-5 math measures to the Common Core Standards (Technical Report # 1229)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Park, B. J., Irvin, P. S., Anderson, D., Alonzo, J., & Tindal, G. (2011). *A cross-validation of easyCBM reading cut scores in Oregon: 2009-2010 (Technical Report # 1108)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Sáez, L., Irvin, P. S., Alonzo, J., & Tindal, G. (2012). *Phoneme segmenting alignment with the Common Core Foundational Skills Standard Two: Grades K-1 (Technical Report # 1227)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Sáez, L., Irvin, P. S., Alonzo, J., & Tindal, G. (2013). *Alignment with the Common Core State Standards: easyCBM K-3 word reading (Technical Report # 1303)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Sáez, L., Park, B. J., Nese, J. F. T., Jamgochian, E. M., Lai, C. F., Anderson, D., Kamata, A., Alonzo, J., & Tindal, G. (2010). *Technical adequacy of the easyCBM reading measures (Grades 3-7), 2009-2010 version (Technical Report # 1005)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Saez, L., Whitney, M., Swanson, D., & ALonzo, J. (2021). *The alignment between easyCBM© mathematics and literacy assessments and state and national standards (Technical Report # 2101)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Saven, J. L. (2014). *National norms*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Saven, J. L., Tindal, G., Irvin, P. S., Farley, D., & Alonzo, J. (2014). *easyCBM norms 2014 edition. (Technical Report # 1409)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Tindal, G. (2017). *Oral reading fluency: Outcomes from 30 years of research (Technical Report # 1701)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Tindal, G., Nese, J. F., & Alonzo, J. (2009a). *Criterion-related evidence using easyCBM® reading measures and student demographics to predict state test performance in Grades 3-8 (Technical Report # 0910)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Tindal, G., Nese, J. F., & Alonzo, J. (2009b). *Hierarchical linear modeling of passage reading fluency growth as a function of student characteristics (Technical Report # 0922)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Tindal, G., & Nese, J. F. T. (2024). *Criterion validity and classification accuracy of easyCBM: Grades 3-8 (Techncial Report # 2401)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Discussion

Research on *easyCBM*© is extensive and research on CBM in general is even more extensive. We conclude with reference to three publications that are highlighted for their comprehensiveness.

- Tindal (2013) provides a comprehensive summary of research on CBM in reading, writing, and mathematics. This review reaches back to the initial research behind the measures used for easyCBM©, which is a direct descendent of this research.
- Anderson et al. (2014) summarize extensive research on all aspects of *easyCBM*© including test development, reliability, and decision-making. They summarize major findings (until that publication date) on all measures of reading and mathematics.
- Carlson, Geisinger, and Jonson (2017) published summaries of *easyCBM*© in the *Buros Mental Measurement Yearbook* by Brookhart and Hawley (pages 317-320). This series may be the most prestigious source of test reviews in the field of both education and psychology and provide information on extensive technical adequacy.

In the end, these reviews and others on the technical adequacy of *easyCBM*© provide educators guidance for interpreting student performance with a reference: norm-referenced, individual-referenced, and criterion-referenced.

- **Norm-referenced** measures compare student performance to peers. Established norms help identify students at risk using percentile ranks. These ranks show performance relative to others at the same grade level.
- **Individual-referenced** measures track student progress over time. Frequent assessments allow teachers to document changes in performance, make instructional adjustments, and compare current performance with previous results for the same student.
- **Criterion-referenced** measures target specific skills. Teachers use item-level reports to identify areas of difficulty, focusing instruction on skills with which students struggle. Mastery is tracked and rewarded through levels.

easyCBM©'s references support risk identification, progress monitoring, and diagnostic decision-making, enhancing student performance interpretation and educational outcomes.

Discussion References

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *The Standards for Educational and Psychological Testing.* American Educational Research Association.

Anderson, D., Alonzo, J., Tindal, G., Farley, D., Irvin, P. S., Lai, C. F., Saven, J. L., & Wray, K. A. (2014). *Technical manual: easyCBM (Technical Report # 1408)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Carlson, J. F., Geisinger, K. F., & Jonson, J. L. (2017). *The Twentieth Mental Measurements Yearbook*. Buros Center for Testing, University of Nebraska.

Swanson, D., & Tindal, G. (2024). *An authoritative bibliography of technical adequacy research conducted on easyCBM – 2024 (Technical Report # 2402.2)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Tindal, G. (2013). Curriculum-based measurement: A brief history of nearly everything from the 1970s to the present. *International Scholarly Research Network (ISRN Education)*, *2013*(Article ID 958530), 29 pages. http://dx.doi.org/10.1155/2013/958530