The University of Iowa

ITP

Iowa Testing Programs

THE IOWA TESTS®

Vertical Scaling and the Assessment of Growth

What is Vertical Scaling?

Vertical scaling is the term used for the process of linking assessments to describe student growth over time. Although the methods can be complex, the goal is quite simple: to create a framework and metric for reporting the educational development of individuals and groups. The challenge of vertical scaling of assessments has existed since the first use of standards-based assessments to measure individual and group progress (Patz, 2007). Today, vertical scaling is needed for assessments of growth toward college and career readiness standards and for adaptive testing. In these applications, comparative information about results from assessments of different levels of difficulty is needed to build a vertical scale. When the assessments are aligned to appropriate cross-grade content standards, valid use of test scores as indicators of growth based on standards can be achieved.

Points on a vertical score scale are a kind of cognitive map to the future. They provide a basis for describing progress, setting goals, and ultimately determining whether students are on track for college and career readiness.

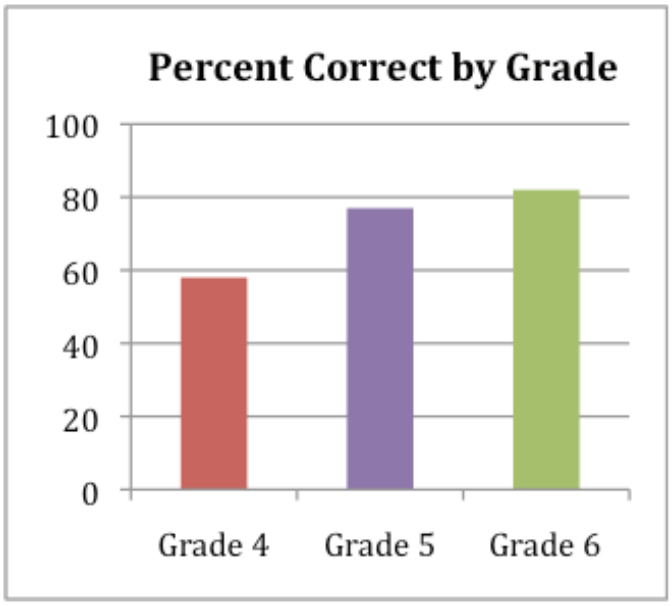
Essential Content for Growth—The Learning Continuum

Developing a scale for growth means defining content standards that describe continuous learning. It also means measuring progress toward those standards with valid and reliable assessments that are true to the goals and objectives implied by the standards. The term "vertical alignment" has been used to convey the message that, to assess growth, changes in test content must reflect a meaningful progression of complexity in what students know and are able to do (Wise & Alt, 2005). Vertically aligned content standards provide specifications for the learning continuum of an achievement domain.

A learning continuum is the foundation of any scope and sequence of instruction and of any assessment program designed to measure growth. In the early stages of learning, students may make rapid progress with respect to relatively concrete aspects of domain knowledge and cognition, whereas in later stages higher levels of cognition and greater content complexity characterize continued progress. The five stages of development in reading (Chall, 1996) are a good example of a learning continuum – there is an underlying construct and a progression that describes how children change from "learning to read" to "reading to learn." In this sense, the learning continuum constitutes a broad definition of the achievement domain and what it means to "grow" with respect to important content standards or guideposts of the domain.

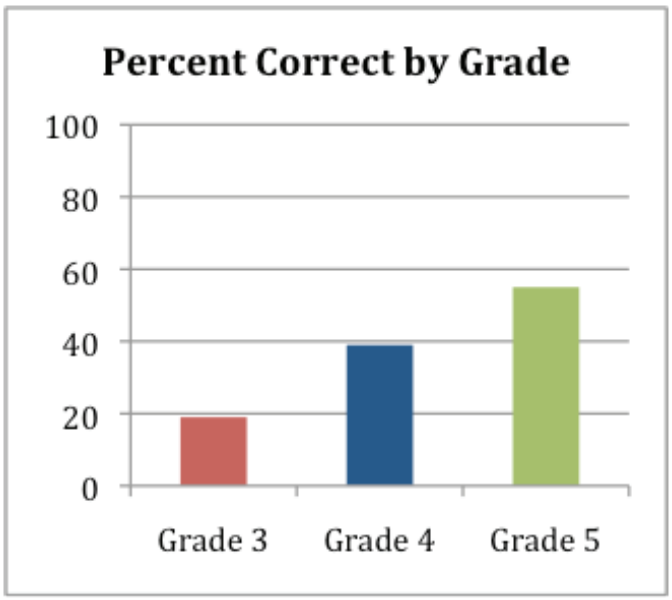
In the line marked with ✓, an insect that is "acrobatic" is

A easily found.
B a skillful flier.
C beautiful to look at.
D prey for other insects.



Four friends decided to meet each other at the park. Kate walked ¼ mile. Terry walked 1¾ miles. Denny walked ¼ mile. Sondra walked 1¾ miles. What is the best estimate of the total miles they walked

A 2
B 4
C 8
D 10



Developing Assessments for Growth—Keeping the Domain in Focus

Assessing a child's location on the learning continuum involves the development of measures aligned to broad content standards and reflecting a level of cognitive complexity appropriate for that child's stage of development. Developmental appropriateness is (1) guided by research and practice in the achievement domain (e.g. content standards in consensus documents of national organizations such as the National Council of Teachers of Mathematics and the International Reading Association; curriculum guides and standards adopted by state departments of education; and, most recently, the Common Core State Standards) and (2) established through extensive field testing of assessment materials in *multiple grades*. Valid and reliable measurement of growth requires both.

A few concrete examples illustrate the characteristics of test items that contribute to the assessment of growth and the research evidence necessary to establish those characteristics. The sidebar shows two items included in recent field test research for the new Iowa Tests and the percent of students who answered each item correctly in three adjacent grades.

The first item (shown without the accompanying reading passage) is aligned to Common Core ELA Standard Vocabulary Acquisition and Use: use context as a clue to the meaning of a word or phrase. It shows a marked increase from 58 percent correct to 77 percent correct between grades 4 and 5, but only a small increase between grades 5 and 6 to 82 percent. The second item is aligned to Common Core Math Standard Build Fractions from Unit Fractions: add and subtract mixed numbers with like denominators. It shows marked increases in percents correct from 19 to 39 to 55 in grades 3, 4, and 5, respectively. Both of these items provide evidence of measureable growth with respect to age and grade appropriate content from the Common Core State Standards. The reading item supports measured growth best between grades 4 and 5, whereas the math item supports growth equally well across all three grades in the research study.

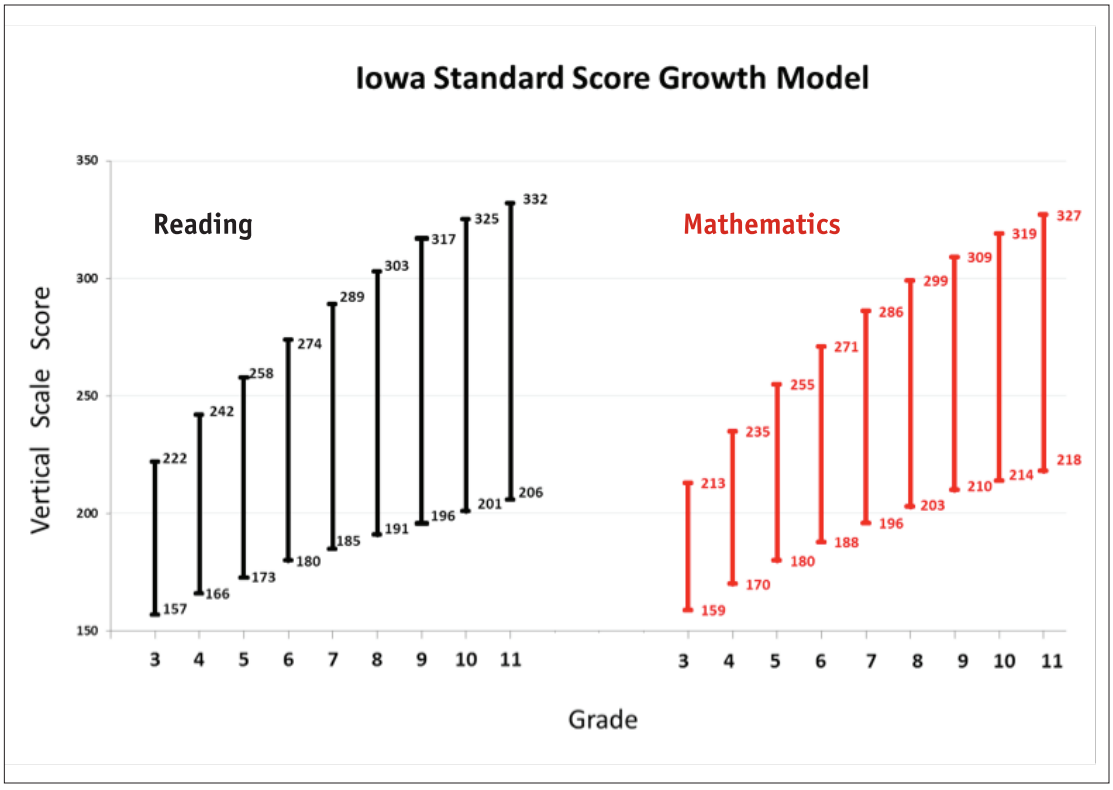
How can growth information be used?

Appropriate use of this information with other available evidence about individuals and groups can lead to instructional planning and curriculum changes (e.g. in a Response-to-Intervention or RTI framework). This information can be used to set goals for the upcoming school year that will lead to student growth that exceeds expectations.

Developing an Evidence-Based Vertical Scale

Assembling test forms with an evidence-based approach to growth on established content standards is a key element in vertical scaling. The methods used to build a vertical scale will only work as intended if the assessment being scaled yields meaningful and stable changes in achievement across time. Assessments matched to content that is not vertically aligned across grade or that reflects an overly granular approach to domain definitions and content specifications may show irregular patterns of growth across grades for both individuals and groups.

The conceptual framework for a vertical scale is established when the content standards and learning progressions of the achievement domain are determined. In developing vertical scales for The Iowa Tests, special assessments were designed in each content area based on the standards. These assessments were wide-range achievement tests consisting of items that spanned multiple grade levels to provide comparative information about the expected performance of students at different developmental levels of the content-area learning continuum. The items presented previously would be examples of items that could appear on these special assessments as well as on grade-level operational forms. Comparative results on the special assessments administered across grades were used to define the range of student performance within each grade level and the amount of overlap between the distributions of student scores at different grade levels. Finally, a numerical scale that describes the growth pattern observed on the special assessments was determined, resulting in what is called the Iowa Standard Score Growth Model. The growth model for Mathematics and Reading in grades 3 through 11 is shown in the figure below.

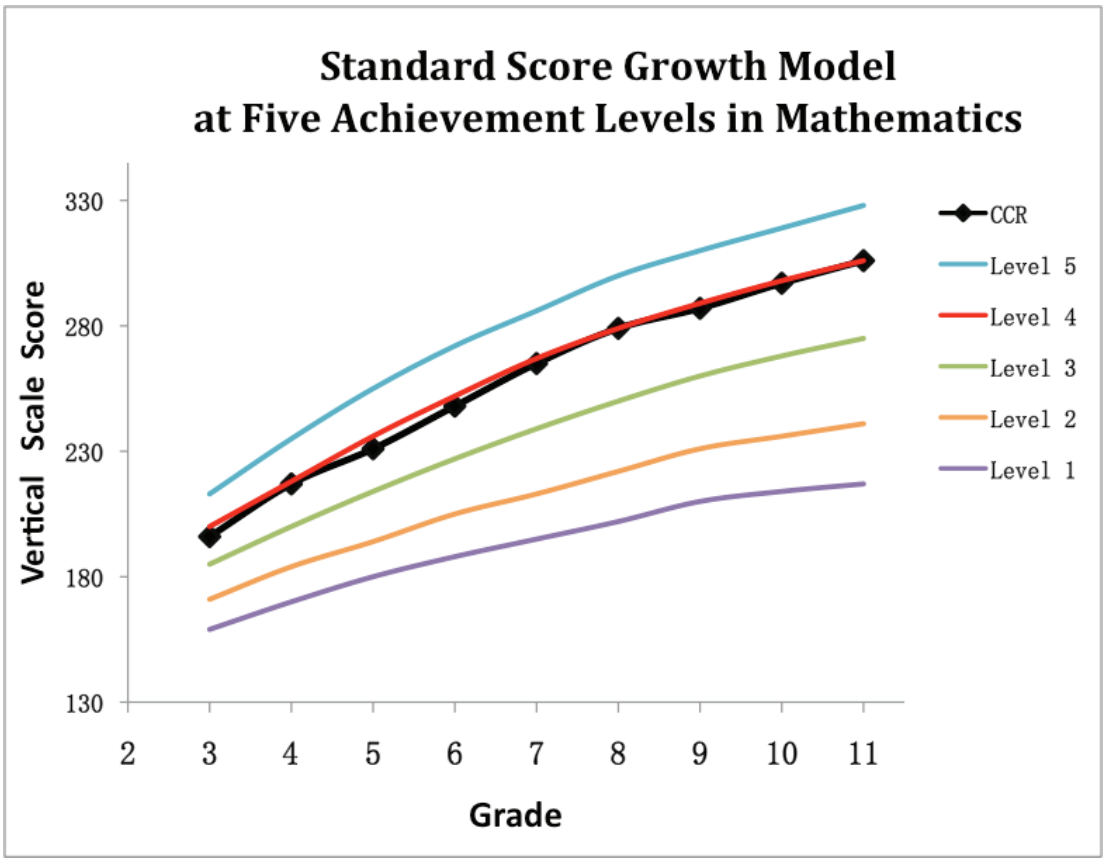


Developing an Evidence-Based Vertical Scale (continued)

Several important characteristics of the Iowa Growth Model are evident in the vertical scale plots for Mathematics and Reading given in the figure. First, the range of vertical scale scores increases from grade 3 to grade 11. This is a consequence of increasing variability observed on the special assessments used in scale development—student achievement becomes more variable over time. Second, the scale ranges overlap substantially in adjacent grades and considerably across grade spans of 3 to 4 years. In terms of location on the learning continuum in Reading, for example, many students in grade 5 are expected to exceed students in grade 6 in comprehending what they read. Fewer are expected to exceed the comprehension of students in grades 8 or 9.

Because evidence about overlap between grades was used in vertical scale development, the resulting comparative information about performance across grades creates a research-based frame of reference for the location of individuals and groups on the learning continuum. That same evidence is the basis for determining expectations for growth in subsequent years from a child's current achievement level and Iowa Standard Score. It can also be used to develop student trajectories to support "on track" interpretations related to college-and-career readiness as long-term education goals.

Another characteristic of the Iowa Growth Model and Standard Score Scale is illustrated in the figure below, which shows expected growth in mathematics for students at various levels of achievement. Depending on a student's current achievement level, the growth model provides an expected Iowa Standard Score for subsequent years. The amount of expected growth increases with achievement level. This trend is reflected in longitudinal analyses of achievement at local, state, and national levels. Importantly, it describes expected growth. In addition, information about student growth can be placed in a college and career readiness context (CCR—dark trend line) to support use of results by school counselors and parents interested in long-term goals and opportunities for postsecondary education.



References

Chall, J.S. (1996). *Learning to read: The great debate.* Fort Worth: Harcourt Brace College Publishers.

Huynh, H. and Scheider, C. (2005). Vertically moderated standards: background, assumptions, and practices. *Applied Measurement in Education. 18*(1), 99–113.

Patz, R. (2007). *Vertical scaling in standards-based educational assessment and accountability systems*. Washington, DC: Council of Chief State School Officers.

Wise, L. and Alt, M. (2005). Assessing vertical alignment. Washington, DC: Council of Chief State School Officers.

The University of Iowa
ITP
Iowa Testing Programs

Iowa Testing Programs
College of Education
University of Iowa
Iowa City IA 52242-1529

steve-dunbar@uiowa.edu
catherine-welch@uiowa.edu

www.education.uiowa.edu/itp