

HIGHER EDUCATION UPDATE

NUMBER UP/01-2
FEBRUARY 2001



News from the

CALIFORNIA POSTSECONDARY EDUCATION COMMISSION

Alan S. Arkatov, *Chair*
Carol Chandler, *Vice Chair*
Phillip J. Forhan
Robert A. Hanff
Lance Izumi
Kyo "Paul" Jhin
Velma Montoya
Ralph R. Pesqueira
Guillermo Rodriguez, Jr.
Evonne Schulze
Khyl Smeby
Howard Welinsky
Melinda G. Wilson

Warren H. Fox
Executive Director

1303 J Street, Suite 500
Sacramento, California 95814-2938
Telephone (916) 445-7933 (Voice)
FAX Number (916) 327-4417

Assessment and Change: The Role of Student Testing in California Education

A high-stakes enterprise

Student testing has enormous implications for eligibility, remediation, educational equity, and educational opportunity for all students. The California Postsecondary Education Commission has long had an interest and involvement in these issues. This document provides both an overview of the history and current uses of student testing in California and a brief summary of some of the issues in the debate over assessment.

In addition to their utility in assessing knowledge and abilities of individuals, test data provides educators and policy makers with information about the aggregate performance of students from which inferences can be made about the overall quality of teaching and educational services in a given school setting. Although there are many ways to evaluate student achievement, standardized tests are often chosen because of their reliability and means of providing an objective assessment of student performance. Consistent administration and scoring procedures enable evaluators to interpret scores in a uniform manner and compare results of different test takers.

In recent years, student testing has become increasingly a high-stakes enterprise in California. Concerns about low student achievement scores have prompted calls for tougher standards and a demand for higher quality schools has led to greater reliance on test data as a gauge of student progress and institutional accountability.

The publication of *A Nation at Risk* (National Commission on Excellence in Education, 1983) is often cited as a catalyst in the new nationwide movement toward greater public school accountability. The reported low student-achievement scores alarmed parents, educators, and policy makers, and elicited a call for tougher academic standards and stronger, better schools. As a result, the measurement of school quality shifted from resource tabulation to student learning assessment. Policy makers began to look at standardized tests as a means of measuring student progress and evaluating educational quality.

Today, policy makers are looking anew to various forms of testing to measure how well students, teachers, and schools are meeting these higher, tougher expectations. California's testing policies offer both a "carrot-and-stick" approach to accountability. Teachers and schools receive incentive funding when student test scores go up, but are labeled "low performing" for below average test scores.

These high stakes extend to the world well beyond school. The technology-driven economy of the 21st century will require a workforce with advanced skills, creating very real "high stakes" for California's economy. Many argue that, in order to remain competitive in the new economic environment,

California must ensure that future workers possess the knowledge and skills needed to be productive citizens and successful individuals.

California higher education institutions, also concerned with demonstrating effective teaching and learning, share with K-12 a concern for ensuring a more natural and seamless connection between the testing efforts of K-12 and higher education admission and student success.

History of student testing

Student testing has long been a part of schools and education. During the Middle Ages, oral examinations were common for some disciplines in European universities. Oral examinations for the Bachelor's and Master's degree were first used at Oxford in 1636 (Popham, 1990). In early America's one-room schoolhouses, tests were primarily oral and were probably used to determine if students were ready to advance to more challenging studies. In their most basic form, tests provide a way for students to demonstrate what they know and for teachers to gauge student learning.

American schools first tested students in a consistent, standardized manner in 1845 when the Boston school system began using preprinted, short-answer tests. Such early efforts helped educators understand the complexity of the teaching and learning relationship and the demonstration of knowledge.

Pioneer efforts in modern testing

Pioneer efforts of more modern testing began at the turn of the century, with Joseph Mayer Rice and Edward Thorndike among the first to use standardized tests to establish average scores for different grade levels. Thorndike is believed to be the first to establish achievement norms for a variety of subjects, including arithmetic, reading, and handwriting. By administering a test to a large number of students and summarizing the average performance of the "norm group," Thorndike's norm-referenced methodology allowed educators, for the first time, to compare their students' performance with the average performance of others.

As other pioneers emerged, so did new forms of standardized tests that measured innate ability and allowed educators to "predict" future performance. Among them, French psychologist Alfred Binet created a scale in 1905 he claimed could ascertain whether a child would benefit

from instruction. By being able to estimate a child's "mental age," the Binet test claimed it could identify a student's capacity to learn. A Stanford University psychologist, Lewis M. Terman, took Binet's work a step further by dividing the "mental age" by the chronological age to arrive at what Terman called the Intelligence Quotient, or so-called "IQ."

The Stanford-Binet scale

The resulting Stanford-Binet scale, published in 1916, provided instruction on the use of the Binet test to assess a student's inherent intelligence and capacity to learn. The first wide-scale use of intelligence tests to determine individual aptitude came at the outset of World War I. Looking for a way to identify potential officer candidates and make more effective staffing decisions, the U.S. Army administered multiple-choice intelligence tests to more than 1.7 million men. The success of the Army's testing program led to a proliferation of standardized, multiple-choice item aptitude tests and, by 1922, these IQ tests were used frequently by educators to classify and track school children.

Achievement tests were later developed to gauge student performance relative to a standard curriculum. The Stanford Achievement Test, for example, was developed to measure and compare student achievement in various basic subjects, including reading, mathematics, and science. Based on national curriculum standards and instructional methods, and indexed to a large national sample of students, the exam became the model for many standardized tests. The Stanford Achievement Test, Ninth Edition, (Stanford 9), which is used in California's current Standardized Testing and Assessment Reporting program, is based on a national standardization sample of more than 500,000 students.

Achievement tests have also been used by some colleges to form admission decisions. The interest of James Conant, Harvard University President from 1933-1953, in establishing an objective method of evaluating applicants and selecting scholarship students led to the use of the Scholastic Aptitude Test (SAT) in college admission decisions. Now called the Scholastic Achievement Test, the SAT is an integral part of the college and university application process and an annual rite of passage for millions of students. However, the need to assess mastery of specific subject areas rather than aptitude is provoking discussions about reducing the role standardized tests should play in admission decisions.

Types of tests

The major categories of tests are norm-referenced and criterion-referenced. While both may be “standardized,” the main differences are in the intended purpose of each and in the way test scores are interpreted.

Norm-referenced tests are the most common and are used to compare student performance with a “norm” or average performance of a sample peer group. The content and criteria may or may not reflect any particular curriculum or academic goals established by governing boards or policy makers. Because norm-referenced tests are typically “normed” against large samples of test takers, their content usually reflects generalized knowledge or skill sets. Examples of norm-referenced tests used presently in California public schools include the SAT I and the Stanford 9.

Scores from norm-referenced tests are interpreted by indexing individual performance relative to the average performance of students in a comparison group, and are often expressed in terms of standing in a percentile group. For example, a score of 1220 on the 1998 SAT I was in the 75th percentile, meaning that the student’s performance was better than 75 percent of the students in the norm group.

Norm-referenced test data lends itself to sorting or rank ordering students and often provides an objective basis for determining and addressing a student’s educational needs. Top scoring students might be selected for enrichment programs, like California’s Gifted and Talented Education (GATE) program, while those scoring below 50 percent might receive special education or extra help.

While scores from a norm-referenced test provide information about how a student performs relative to others, it may or may not show individual student learning relative to specified achievement standards. Criterion-referenced tests, in contrast, provide information about how individuals perform relative to defined criteria. These tests help educators ascertain what test takers can do and what they know. Because such tests provide information about how well students perform relative to defined learning outcomes, they are useful in determining the progress students are making toward mastering the expected knowledge and skill levels. Ideally, the content of criterion-referenced tests reflects the curriculum. When standards, curriculum, and testing are aligned, criterion-referenced

tests provide rich data about the performance of students and teachers.

In the early 1990s, the assessment movement attempted to bridge the gap between these basic test styles. Critics argued that multiple-choice questions, the standard format for many tests, do not assess a student’s ability to come up with his or her own answers. This led to including open-ended questions and less structured, less standardized approaches to gauging student achievement.

Performance-based assessments, which help educators look at how test takers demonstrate competency, and portfolios, which assess growth over time, are two popular alternative assessment methods. These tend to provide a more “complete” picture of student performance, allowing evaluators to understand responses in context. However, these assessments are often expensive to administer and can raise questions about measurement error and bias. The California Learning Assessment System (CLAS), for example, met with public outcry in the late 1990s about test questions and reading material and was ultimately abandoned as a public school testing program.

In order for tests to be effective measurement devices, they must be used for their intended purpose and they must provide reliable and valid measurements. A norm-referenced test is appropriate if the question is, “How do our students compare with other students in basic skill areas?” A criterion-referenced test would be better if asking, “How well are our students learning what we want them to learn?” Whether norm-referenced or criterion-referenced, good tests must be consistent in their ability to measure performance and accurately represent what they claim to measure.

K-12 testing in California

Tests used in California public schools attempt to answer a variety of questions about the performance of students and teachers. California’s new K-12 content standards and public demands for more accountability has helped ensure that testing assumes a prominent role. California uses a variety of norm-referenced, criterion-referenced, and standards-based exams to assess how students compare to others as well as progress in achieving specific learning goals. There are now more than 10 different exams that are used to determine achievement, proficiency, and college eligibility placement. These include:

The Stanford 9/ STAR program

California adopted the Standardized Testing and Assessment Reporting (STAR) program in October 1997 after looking at various approaches to assessment. Enacted by SB 376, it required the California State Board of Education to develop academic content standards and the systematic testing of nearly all public school children. The

| <i>Achievement Tests</i> | <i>Proficiency/Diagnostic/ Placement Tests</i> |
|--------------------------------------|---|
| The STAR Program | Entry Level Mathematics Exam |
| - Stanford 9 | English Placement Test |
| - California Standards Test | TOEFL |
| - Primary Language SABE/2 | Subject A |
| Golden State Examination | AP/IB Exams |
| High School Exit Exam | MDTP |
| English Language Development | |
| | |
| <i>College Entrance Exams</i> | <i>Achievement/Proficiency</i> |
| PSAT | NAEP |
| ACT | Physical Fitness |
| SAT I | |
| SAT II Subject Tests | |

State Board designated the *Stanford Achievement Test* series for the STAR program. The norm-referenced SAT 9 assesses a broad range of basic academic skills and provides comparable individual pupil scores. Students in grades 2 through 8 are tested in reading, mathematics, writing, and spelling. Students in grades 9 through 11 are tested in reading, writing, mathematics, science, and history/social science. The exam was given first in the spring of 1998, and was augmented in 1999 to include questions aligned with recently adopted academic content standards. This standards-based portion of the STAR program is called the California Standards Test.

Limited-English-proficient (LEP) students who have been enrolled in any California school less than 12 months must be tested in their primary language provided a State Board-designated test in that language is available. For students whose primary language is Spanish, the Spanish Assessment of Basic Education (SABE/2) exam is designated. The Board has yet to designate tests for students in other languages.

Reported statewide and by county, district, and school, the results help answer the question about “how well schools carry out instruction.” The Public School Perfor-

mance Accountability Act, established by SB 1X (Alpert, Statutes of 1999), created rewards and interventions for schools and educators as a means of improving student performance. The act also provides an “Immediate Intervention” grant program for helping underperforming schools and an incentive grant program for high achieving/improving schools.

The Golden State Exam Program

The Golden State Examination (GSE) program is a voluntary examination program available to students in grades 7-12 providing rigorous end-of-course achievement examinations in 18 college-preparatory subject areas and provides a means of recognizing students who demonstrate outstanding levels of achievement in several areas. Districts are required to make Golden State Examinations available to all students. High school graduates who attain the three highest levels of achievement designations (high honors, honors, or recognition) on six such examinations are eligible to receive a Golden State Seal Merit Diploma. (Less than 1% of California’s 1998 high school graduates did so.)

Mathematics Diagnostic Testing Program (MDTP)

These tests are administered by high school and middle school teachers and provide diagnostic information to teachers for the purpose of adjusting curriculum and determining student readiness for advancement to subsequent levels of mathematics. The program was developed jointly by California State University (CSU) and University of California (UC) faculty members and consists of five tests: Algebra Readiness, Geometry Readiness, Second Year Algebra readiness, Mathematical Analysis Readiness, and Calculus Readiness. Although the UC has no official system-wide policy on mathematics placement; however, many UC campuses use the Mathematics Diagnostic Testing Program (MDTP) to make placement decisions about mathematics.

Assessments in Career Education

Similar to the GSE program, the Assessments in Career Education (ACE) is a voluntary, end-of-course testing program that recognizes outstanding levels of student achievement in five career-technical areas: agriculture core; computer science and information services; health care; food services and hospitality; and technology. Test

results provide potential employers with a record of a student's accomplishments in preparing for technical careers in the categories listed.

Physical Fitness Testing

This criterion-referenced testing program is required by Education Code Section 60800. All students in grades 5, 7, and 9 are required to take a physical fitness test each spring. The testing data is used to track the development of fitness programs and compare student performance in to national norms. Legislation passed in 1998 requires the California Department of Education to compile the performance test results and submit annual reports to the Governor and the Legislature by December 31 of each year.

National Assessments of Educational Progress

The National Assessment of Educational Progress (NAEP) is administered by the U.S. Department of Education. For over 30 years, this criterion-referenced test has provided information on the educational achievement of American students in grades 4, 8 and 12. Assessments are conducted periodically in various subjects including reading, writing, mathematics, science, history, geography, and other fields. The results include student performance data and instructional factors related to that performance. The 2000 NAEP exam included mathematics and science. In 2001, NAEP will administer national assessments in history and world geography to a sample of California's school children.

High School Exit Exam

California is in the process of establishing higher standards for high school graduation and developing a High School Exit Examination (HSEE). The purpose of the exam is to "ensure that pupils who graduate from high school can demonstrate grade level competency in reading, writing, and mathematics."

Senate Bill 2X (1999) requires that, beginning in the 2003-2004 school year, students will not receive a high school diploma without first achieving a passing score on an exit examination. Beginning in 2000-01, 9th grade students have the option of taking the exam. In 2001-02, 10th grade students will be required to take and pass both the writing and mathematics sections of the exam. These students may take the exam during each administration

until they pass each section. The exam is to be aligned with California's academic content standards for language arts and mathematics.

Higher education assessments

High school students not only are subject to the achievement tests listed above, but many also take many tests for purposes of pursuing higher education. In general, these tests serve as either a tool for admissions purposes or for placement. Below is a brief description of some of the major tests used by higher education institutions in California.

Scholastic Assessment Test I (SAT I)

The SAT I is a norm-referenced, multiple-choice exam that measures mathematical and verbal reasoning. It is used for college selection and is intended to predict success in college. Currently, both UC and CSU utilize the SAT I. All applicants to UC are required to submit SAT I scores, although a student may choose to take the ACT. Applicants to CSU with less than a 3.0 high school grade point average must submit their SAT or ACT scores for admission.

Scholastic Assessment Test II (SAT II)

The SAT II is a set of one-hour, norm-referenced, multiple choice and open-ended tests used primarily for admission to the more selective colleges and universities. Proponents suggest such tests are more closely aligned to the academic content in particular subject areas and therefore better predictors of student achievement. The University of California requires that to be eligible for admission, all applicants must take three SAT II tests. In addition, the University has recently begun assigning greater weight to the SAT II tests than it had previously in determining admission.

The American College Testing (ACT)

This exam is a three-hour, multiple-choice exam that assesses achievement in several academic subjects. It is used primarily used for college admission. Although the vast majority of students in California take the SAT rather than the ACT, approximately 38,000 California students took the ACT in 2000. Both CSU and UC accept the ACT as well as the SAT.

Test of English as a Foreign Language (TOEFL)

This test is for students who have not attended schools at the secondary level or above for at least three years full time where English is the principal language of instruction. It is intended to measure English proficiency.

Advanced Placement (AP) Exams

These exams are used to measure college-level achievement in many different subject areas and to award academic credit to high school students who demonstrate college-level proficiency. Scores ranging from 1 (No recommendation) to 5 (Extremely well qualified) are awarded to students who take the exams. These scores are used as evidence of the students' abilities and achievements and allow colleges and universities to make their decisions regarding whether or not to grant credit and/or advanced placement.

Entry Level Math (ELM) Exam

This examination is designed to assess the skill levels of entering CSU students in the area of mathematics typically covered in the three years of rigorous college preparatory mathematics courses in high school. Those undergraduate students who do not demonstrate college-level skills will be directed to courses or programs designed to help them attain these skills.

Most entering undergraduates at CSU take the ELM examination before enrolling in a course that satisfies the college-level mathematics requirement of the General Education-Breadth program. The exceptions are those who score sufficiently high on the mathematics section of the SAT I: Reasoning Test, SAT II: Mathematics Test, ACT Mathematics Test, Advanced Placement Mathematics Examination (Calculus AB or BC), Advanced Placement Statistics Examination, or transfer students who have satisfied the Quantitative Reasoning requirement under the Intersegmental General Education Transfer Curriculum.

English Placement Test (EPT)

This test, used by the CSU system, is designed to assess the level of reading and writing skills of entering undergraduate students so that they can be placed in appropriate baccalaureate courses. The CSU English Placement Test must be completed by all entering undergraduates with the exception of those who score sufficiently high on

the SAT Reasoning Test, the SAT II Writing Test, AP Language and Composition or Literature Exam, or transfer students who have completed the English Composition requirement under the Intersegmental General Education Transfer Curriculum.

Subject A Exam

The University of California requires entering students who do not meet the UC standards for English to take a two-hour essay exam. Students are required to read a passage and then write an essay responding to a single topic based on the content of the passage.

All students who enter the UC directly from a California high school must take the University-wide Subject A Examination unless they have a score of 680 on the SAT II Writing Test or a 3 or higher on the Advanced Placement Examination in English. Students may take the exam only once and those who do not pass can fulfill the requirement by either achieving a satisfactory SAT II Writing or AP English score, complete with a "C" grade or better an acceptable college course in English composition, or achieve a score of 5 or better on the International Baccalaureate's Higher Level English A Examination. Students who have not satisfied the Subject A requirement prior to enrolling in the University must take and pass with a "C" grade or better, a writing course designated by their campus for satisfying the Subject A requirement.

Diagnostic Writing Services (DWS)

This service is available for students and teachers to assess current writing skills against college-level expectations. It allows individual students to write an essay in response to an actual English Placement Test essay prompt and to submit the essay via the Internet. A university EPT reader uses the scoring rubric of the EPT test and assesses student writing against the standards expected of entering college students. Diagnostic statements are provided to the student via the Internet.

Writing Proficiency Exam

All candidates for a bachelor's degree at CSU must take and pass, prior to graduation, a test that assesses a student's ability to write with college-level proficiency. The student must compose an essay from a topic that is provided to them.

Current efforts

The California Education Roundtable -- which includes the Chief Executive Officers from California Community Colleges, California State University, University of California, the Postsecondary Education Commission, and Superintendent of Public Instruction -- has long recognized the importance of collaboration and cooperation between K-12 and postsecondary education to ensure that assessment of students throughout the educational continuum is better aligned and streamlined.

To that end, the Education Roundtable has convened a working group to begin to develop policy principles on the issue of alignment and to formulate a more logical and articulated testing and assessment system for California's students. The Commission is participating in these meetings.

Conclusion

As policy makers continue to consider ways to improve California's educational system, strengthen the quality of instruction provided to our students, and raise achieve-

ment levels, there will likely be many initiatives aimed at measuring the extent to which we are in fact meeting the new standards and higher expectations that have been established for students. Used appropriately, testing can be a useful tool for demonstrating improvement, measuring the quality of instruction, and determining gains in student achievement across the educational continuum.

California has invested vast resources in recent years to creating uniformity in the content of our curriculum, reducing class sizes, strengthening teacher quality, and setting higher standards for educators and their students. It is important that the process selected to evaluate these efforts informs the State's ability to continue the course toward better educational institutions -- elementary through university -- stronger instruction, and successful students.

The California Postsecondary Education Commission will continue to monitor the activities underway in California on the issue of testing, participate in discussions intended to enhance the alignment of K-12 and higher education assessments, explore the policy implications associated with the topic, and provide recommendations for Commission consideration, where appropriate.