

Requirement 4

A state must provide educators with optional classroom assessment procedures that can measure students' progress in attaining content standards not assessed by state tests.

Because —

- *content standards that are not assessed by state tests are important and should be given instructional attention;*
- *educators need good assessment tools to monitor students' achievement and rarely have the time and resources to develop such tools; and*
- *assessments that are routinely administered by educators can and should be used to provide a complete picture of what students know and are able to do.*

For purposes of per-standard assessment, statewide tests must measure a limited number of high-priority content standards. Yet, the Commission believes that instructional and assessment attention also must be given to the remaining state content standards.

Requirement 4 obliges states to develop optional classroom assessments for content standards not assessed by state tests. These assessments will support educators' efforts to teach a wide range of skills and knowledge.

We wish to stress that these classroom assessments are *optional*, not mandatory.

Although existing federal guidelines require states to assess all essential content standards, educators must be free to use or develop classroom assessments other than those provided by the state. We believe, however, that states have a responsibility to see that educators understand optional state-developed classroom assessments and know how they could benefit students. We recommend that states conduct professional development activities for educators regarding how to best use these optional assessments for instructional improvement. And we suggest that this professional development also focus on how educators can design their own classroom assessments to measure students'

Requirement 5

A state must monitor the breadth of the curriculum to ensure that instructional attention is given to all content standards and subject areas, including those that are not assessed by state tests.

Because —

- *students benefit from a rich and deep curriculum; and*
- *state tests that measure high-priority content standards could narrow curricular coverage unless steps are taken to forestall such narrowing.*

The Commission believes that states need to support educators' broad curricular coverage, especially if state tests focus on a small number of high-priority content standards (See Requirement 1). Because students, schools, and school districts are accountable for performance on state tests, these assessments are powerful motivators. They can influence educators' day to day instructional activities dramatically, and they can inadvertently lead to a narrowing of the curriculum as educators work to ensure that students perform well on state tests. A narrowed curriculum would deal almost exclusively with content assessed on state tests.

We believe states, school districts, and schools must monitor the breadth of the curriculum that students experience to ensure that it includes more than the content assessed on state tests. As noted earlier, we recognize that current federal guidelines require states to teach and assess all essential state content standards, and that assessments of such standards can be carried out at the state, district, and/or school levels. Therefore, we recommend monitoring of curricular breadth at these same three levels using quantitative and/or qualitative methods that states and school districts develop.

Requirement 6

A state must ensure that all students have the opportunity to demonstrate their achievement of state standards; consequently, it must provide well-designed assessments appropriate for a broad range of students, with accommodations and alternate methods of assessment available for students who need them.

Because –

- *all students must be given appropriate opportunities to demonstrate the degree to which they have mastered state content standards; and*
- *federal statutes require that such opportunities be offered in particular ways.*

The Commission believes that state tests must be designed and developed to allow participation of the widest range of students, and thus permit valid inferences about the performance of all students, including those with disabilities, those with limited English proficiency, and those with other special needs.

Requirement 6 obligates states to design statewide assessments or appropriate alternatives (for example, panel review, performance testing, or portfolio assessment) that provide accurate and useful information to teachers concerning the degree to which students with special needs have demonstrated the skills and knowledge described in the state's content standards. This requirement is also consistent with federal laws that obligate states to develop guidelines for school districts about how all students participate, that is, (a) in the general assessment without accommodations, (b) in the general assessment with accommodations, or (c) in an alternate assessment aligned with the state content standards.

Requirement 7

A state must generally allow test developers a minimum of three years to produce statewide tests that satisfy the *Standards for Educational and Psychological Testing* and similar test-quality guidelines.*

Because —

- *tests developed too hurriedly neither support instruction nor supply accurate evaluative information for accountability programs; and*
- *there is a widespread misunderstanding that high-quality achievement tests can be developed in two years or less.*

The Commission's members understand that state policymakers who are eager to improve education may call for state tests to be developed "as soon as possible." We recognize that the authorization for many states' test-development efforts often is dictated by annual or biennial legislative and budgetary cycles. We believe however, that the likely educational impact of these tests is far too important for the tests to be developed improperly, and proper test development takes time. Experience informs us that a minimum of three years is needed to develop a state test to assess high-priority content standards in a way that promotes instructional improvement.

To be more specific, it takes at least three years to (a) prioritize a state's content standards to identify the standards that state tests will assess, (b) analyze the skills and knowledge students must demonstrate for each content standard, (c) develop sufficient numbers of test items for each high-priority standard, (d) evaluate the test items through small-scale pilot tryouts or other review procedures, (e) formally field-test all the test items, and (f) assemble operational test-forms. Moreover, we believe that the kinds of evidence

* See American Educational Research Association. 1999. *Standards for Educational and Psychological Testing*, Washington, D.C.: Author. See also such documents as Pellegrino, J.W., N. Chudowsky, and R. Glaser (Eds.) 2001. *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, D.C.: National Academy Press.

regarding test quality called for in the *Standards for Educational and Psychological Testing* are genuinely important. It takes time, for example, to assemble sufficient and compelling evidence about the validity of a test.

The development of state tests also takes time because all potential contributors, for instance, teachers, administrators, and citizens, must be given meaningful opportunities to participate in appropriate stages of the test-development process.

Requirement 8

A state must ensure that educators receive professional development focused on how to optimize children’s learning based on the results of instructionally supportive assessments.

Because —

- *most educators are unfamiliar with the instructionally supportive assessment system the Commission advocates; and*
- *it is imperative that educators become adept at using such an assessment system for instructional purposes.*

The Commission realizes that the assessment system we recommend is a new approach that will be unfamiliar to many educators. The kind of state tests we describe and the role we have set forth for classroom assessments (see Requirement 4) in the state’s accountability system are new. Educators will need repeated opportunities to learn how best to use the instructional information that state tests will provide. Educators also will need opportunities to either use state-provided optional classroom assessments or to develop their own classroom assessments that provide information about students’ learning. We urge state policymakers to provide educators with a variety of professional development activities designed to increase their success in using the type of instructionally supportive assessment system we advocate.

Requirement 9

A state should secure evidence that supports the ongoing improvement of its state assessments to ensure those assessments are (a) appropriate for the accountability purposes for which they are used, (b) appropriate for determining whether students have attained state standards, (c) appropriate for enhancing instruction, and (d) not the cause of negative consequences.

Because —

- *any assessment system, no matter how well designed, can be improved; and*
- *a state's assessment system must perform both an accountability function and an instructional improvement function.*

The Commission believes that assessment systems need to be evaluated and improved on a continuing basis to ensure that they fulfill their intended purposes. We think this final requirement is especially important because we are recommending a new approach to state testing. The tests the Commission advances in this report represent a new generation of state tests that provide information for accountability purposes as well as information for improving instruction. Furthermore, we have described a new kind of assessment system in which results from state tests that focus on high-priority content standards are combined with results from classroom assessments that focus on other state content standards and, together, provide a more complete picture of students' learning than would be provided by a single form of assessment.

We recommend that states undertake independent evaluations to see if state and classroom assessments function as we have described in this report. We also think it is important for independent evaluations and studies to determine the degree to which a

state's students have had an opportunity to learn what is necessary to attain the state's content standards. And it is important for such studies to provide empirical evidence to demonstrate that assessments specifically designed to detect the impact of effective instruction are, in fact, sensitive to well-conceived instructional interventions. The Commission's chief point here, of course, is that even assessments intended to be instructionally supportive must be regularly scrutinized to make certain that children's learning has been bettered by the implementation of those assessments. Thus, evaluative attention must certainly be given to any unanticipated negative consequences of a state assessment program such as dramatically increased student dropout rates.

A Shared Responsibility

The members of this Commission have set forth requirements we believe must be satisfied for a statewide assessment system to benefit all students. We believe that our nine requirements, if satisfied, will lead to state assessment programs that greatly increase the likelihood of educators' promoting the kinds of achievement a state desires for its students. We also think that these assessment programs will provide policymakers with the evaluative evidence they need to make accountability programs function in the long-term best interests of their state's children.

We have described the assessment system we envision as *responsible assessments for the improvement of learning*. If a state's educational policymakers have an opportunity to install such instructionally supportive tests, yet do not do so, then it seems to us that such a failure to capitalize on an opportunity to benefit children constitutes a form of educational irresponsibility.

We suggest, therefore, that all of those who have an interest in the quality of a state's education system, from its governor to each individual citizen, exert whatever influence they possess to make sure that the assessment system in their state satisfies the Commission's requirements. The education of our children is, after all, a shared responsibility.

Commission Members

Eva L. Baker

A Professor of Educational Psychology and Social Research Methods at the UCLA Graduate School of Education and Information Studies, Eva L. Baker has directed the UCLA Center for the Study of Evaluation (CSE) since 1975. Dr. Baker is also Co-Director of the National Center for Research on Evaluation, Standards, and Student Testing (CRESST), a competitively awarded national institution funded by the U.S. Department of Education. Former president of the Educational Psychology Division of the American Psychological Association, Dr. Baker has been a national officer of the American Educational Research Association, and is a former editor of *Educational Evaluation and Policy Analysis*. A congressionally appointed member of the National Council on Education Standards and Testing, she currently serves on the Independent Review Committee on the Evaluation of Title I Testing. She is co-chair of the committee to revise the standards for Educational and Psychological Testing, and is a member of the Advisory Committee on Education Statistics for the National Center for Educational Statistics. Dr. Baker's research focuses on the integration of instruction and measurement, and she is presently involved in the design of technologically sophisticated testing and evaluation systems of performance assessment in large-scale environments for both military and civilian education.

David C. Berliner

David C. Berliner, Regents' Professor, School of Education at Arizona State University, has taught at the Universities of Arizona and Massachusetts, Teachers College, Stanford University, as well as at universities in Australia and Spain. He most recently served as the Dean of the College of Education, Arizona State University. Dr. Berliner is a member of the National Academy of Education, a Fellow of the Center for Advanced Study in the Behavioral Sciences, and a past president of both the American Educational Research Association (AERA) and the Division of Educational Psychology of the American Psychological Association (APA). From 1970-77, he was Associate Laboratory Director for Research, Far West Laboratory for Educational Research and Development (now WestEd). He is the recipient of awards for distinguished contributions from APA, AERA, and the National Education Association (NEA). Dr. Berliner has co-authored a number of books, including a best-seller, *The Manufactured Crisis*, with B.J. Biddle, *Putting Research to Work*, with Ursula Casanova, and *Educational Psychology* (textbook), with N.L. Gage. He has also authored more than 150 published articles, technical reports and book chapters, and is the co-editor of *The Handbook of Educational Psychology* and the books *Talks to Teachers* and *Perspectives on Instructional Time*.

Carol Camp Yeakey

Carol Camp Yeakey is Professor of Urban Politics and Policy at the Curry School of Education, University of Virginia. She has taught at Columbia University's Teachers College, Rutgers, Purdue University, and De Paul University, where she also served as Vice Provost of Academic Affairs. While at the University of Virginia, she has chaired the Graduate Program in Policy Studies, designed a new graduate degree program in Policy Studies, and established internship experiences in the Washington, D.C. metropolitan region. In addition, she conducts the Annual Policy Institute, which involves national and state legislative leaders. Dr. Camp Yeakey currently teaches graduate courses in education policy research and analysis, advises graduate students, and supervises graduate research. She has been a Ford Fellow of the Academy of Education, a Rockefeller Fellow in Educational Policy Research at Yale, and a Dartmouth Fellow at the College's Center for the Study of Comparative Politics and Intergroup Relations, among other academic honors. Dr. Camp Yeakey is the author of numerous books, articles, and book chapters.

James W. Pellegrino

James W. Pellegrino is Liberal Arts and Sciences Distinguished Professor of Cognitive Psychology and Distinguished Professor of Education at the University of Illinois at Chicago (UIC). He also serves as co-director of a new interdisciplinary center at UIC focused on cognition, instruction, assessment, and teacher development. Prior to coming to Chicago, Dr. Pellegrino was the Frank W. Mayborn Professor of Cognitive Studies at Vanderbilt University, where he also served as co-director of the Learning Technology Center (1989-91) and as Dean of Vanderbilt's Peabody College of Education and Human Development (1992-98). He has taught at the University of Pittsburgh and the University of California at Santa Barbara, where he served as Chair of the Department of Education from 1987 to 1989. For more than 25 years, Dr. Pellegrino has researched children's and adult's thinking and learning, and the implications of cognitive research and theory for assessment and instructional practice. He chaired the National Academy of Science/National Research Council Study Committee for the Evaluation of the National and State Assessments of Educational Progress, which in 1999 produced *Grading the Nation's Report Card: Evaluating NAEP and Transforming the Assessment of Educational Progress*. He is a member of the Board on Testing and Assessment of the National Research Council. Dr. Pellegrino has authored or co-authored more than 185 books, chapters, and journal articles over his career.

W. James Popham (Chair)

W. James Popham is Professor Emeritus, UCLA Graduate School of Education and Information Studies. He has spent most of his career as a teacher, largely at the UCLA Graduate School of Education and Information Studies, where he taught courses in instructional methods for prospective teachers and courses in evaluation and measurement for masters and doctoral candidates for nearly three decades. In January 2000, he was recognized by *UCLA Today* as one of the university's top 20 professors of the twentieth century, and was awarded the 1997 Lifetime Achievement in Educational Research and Measurement by the California Educational Research Association. In 1968, Dr. Popham founded IOX Assessment Associates, a research and development group that created statewide student achievement tests for a dozen states. In 1978, he was elected president of the American Educational Research Association (AERA), and is the founding editor of *Educational Evaluation and Policy Analysis*, AERA's quarterly journal. Dr. Popham has written 20 books, including *Testing! Testing! What Every Parent Should Know About School Tests*, 180 journal articles, 50 research reports, and 150 papers presented before research societies.

Rachel F. Quenemoen

Rachel F. Quenemoen is a Senior Fellow for Technical Assistance and Research at the National Center on Education Outcomes, University of Minnesota. Prior to joining the Center staff in 1999, she spent time as a public health administrator, and she was also a technical assistance coordinator with the Minnesota Department of Children, Families and Learning. She began her career as an educational administrator of regional special and regular education services serving public and nonpublic schools in Minnesota. Ms. Quenemoen has conducted presentations on assessment and accountability issues at forums across the country, and has authored many articles on standards, accountability, and assessment.

Flora V. Rodríguez-Brown

Flora V. Rodríguez-Brown is a Professor of Curriculum and Instruction/Reading, Writing, and Literacy, University of Illinois at Chicago. In addition, she is the Coordinator of Bilingual/ESL Training Programs where she directs the university's recruiting, advising, and training of students in bilingual and ESL education. She also directs a family literacy program and is working on restructuring a public school in Chicago. Dr. Rodríguez-Brown's research includes studies on language proficiency and discourse patterns with bilingual children, attitudes and their relation to second language learning, and the role of transfer and metacognitive strategies in the development of spelling skills in a second language. Since joining the University of Chicago faculty in 1982, Dr. Rodríguez-Brown has taught courses in methodology of language arts and social studies, philosophy of bilingual education, curriculum development for non-

English speaking children. She is the author of numerous articles, book chapters, and monographs.

Paul D. Sandifer

Before becoming a consultant to the Office of Assessment at the South Carolina Department of Education in October, 2000, Paul D. Sandifer was its Interim Director. In that capacity, his responsibilities related to the development, administration, and maintenance of the state's assessment system for grades K-12, as well as the validation of examinations used to certify teachers and administrators. Dr. Sandifer also served as Director of South Carolina's Office of Student Performance Assessment where, among other duties, he was responsible for the state's Basic Skills Assessment Program, including test development, annual administration to approximately 315,000 students, reporting test results, and managing testing contracts. From 1974-1991, Dr. Sandifer was Director of the South Carolina Department of Education's Office of Research. From 1992-1997, he was Assistant Vice President, Development Division, ACT, Inc. (formerly American College Testing). In that capacity, he supervised the development and maintenance of the ACT Assessment Program, the eighth and tenth grade programs, post-secondary placement examinations, and college outcome measures. He has also supervised contracted test development for state education agencies and other organizations.

Stephen G. Sireci

As an Associate Professor in the Center for Education Assessment, School of Education, University of Massachusetts, Amherst, Stephen Sireci teaches graduate courses in test development, statistics, scaling methods, educational assessment, validity theory and research methods, and supervises doctoral students' research. His current research activities include evaluating test comparability across languages, assessing test dimensionality, implementing innovative scaling and standard setting methodologies, appraising test validity, designing computer-based tests and performance assessments, estimating the reliability and validity of scores from complex test designs, improving the attitudes of teachers and minority students towards standardized testing, and refining emerging conceptualizations of validity. As a Senior Psychometrician at the American Council on Education in Washington, D.C., Dr. Sireci directed, supervised, and coordinated research and test development activities related to the Tests of General Educational Development (GED Tests), administered to more than 800,000 adults annually. He also has served as a psychometrician at the American Institute of Certified Public Accountants, and as a pre-doctoral fellow at the Educational Testing Service. Dr. Sireci is the author of many journal articles and book chapters.

Martha L. Thurlow

Martha Thurlow is the Director of the National Center on Educational Outcomes at the University of Minnesota. As Director, she addresses the implications of contemporary U.S. policy and practice for students with disabilities, including national and statewide assessment policies and practices, standards-setting efforts, and graduation requirements. In addition, Dr. Thurlow is a senior research associate, Department of Educational Psychology, College of Education and Human Development, University of Minnesota. She has worked with diverse groups of stakeholders to identify key outcomes for young children and students in grades 4, 8, and 12, and at the post-school level. For the past 25 years, Dr. Thurlow has conducted a wide variety of research studies in special education including assessment and decision making, learning disabilities, and effective classroom instruction. She has authored numerous books and book chapters, and has published more than 200 articles and reports. In 1995, she became co-editor of *Exceptional Children*, the research journal of the Council for Exceptional Children.