

# Forward

**This handbook is organized around criteria for:**

- A. Accountability Reports**
- B. Rewards and Sanctions**
- C. Improvement Requirements and Assistance**
- D. Interventions for Low Performing Schools**
- E. Criteria for Assessment**
- F. Content Standards**
- G. State Level Accountability Reports**
- H. Responsibilities of Other Organizations**

**Each criteria is discussed and examples are provided. The examples are taken from state, districts and other organizations with which NCITE works. NCITE has not reviewed legislation and policies from most states. For some criteria, cautions are discussed under the heading "Challenges to be met and situations to be avoided."**

## **A. Accountability Reports**

### **1. Reports are:**

- a. readily available to the public through media, internet and distribution to parents.**
- b. made clear and meaningful through categorization of levels of school or district performance, such as exemplary, recognized, acceptable and low performing.**

#### **OPTIONS:**

#### **Criteria for placing school in a category are**

- set high when the accountability system is established or**
- set lower at the beginning and then raised systematically over time.**

### **2. Reports include summaries of performance of children disaggregated by:**

- a. income level.**
- b. race.**
- c. primary language.**
- d. performance by quartile.**
- e. continuing students (students who have been at the school for a specified time).**
- f. special education.**

### **3. Reports compare performance to:**

- a. average performance of all students in state or district.**
- b. average performance of students with comparable demographics.**
- c. benchmark performance for students with comparable demographics. Benchmark scores come from the top 10%-15% of the schools in the same demographic group during the previous year.**

### **4. Reports include:**

- a. the percent of students in the school taking the assessments and included in reports.**
- b. percent of students exempted for particular reasons, i.e., special education, language or other criteria.**
- c. analysis of performance on individual topics and standards, if possible.**

# Accountability Reports

## 1. Reports are:

- a. readily available to the public through media, internet and distribution to parents.
- b. made clear and meaningful through categorization of levels of school or district performance, such as exemplary, recognized, acceptable and low performing.

### OPTIONS:

#### Criteria for placing school in a category are

- set high when the accountability system is established or
- set lower at the beginning and then raised systematically over time.

## Example of Criteria 1.a.

Texas Education Code requires wide distribution of state test results to public. (See Texas Education Code 39.053.) Sections of Texas Education Code dealing with assessment and accountability are in State Reports section of this notebook.

**Example of Criteria 1.b.** (See page A-1.) The Texas School Rating System Texas assigns all its public schools an overall rating based on three criteria: dropout rates, attendance rates, and the percentage of students passing each of the reading, writing, and math portions of the Texas Assessment of Academic Skills (TAAS). The state considers the performance not only of the school's entire student body, but also of three racial and economic subgroups. A school cannot earn a rating any higher than called for by its students' weakest performance in any category. Since the rating system was created in 1993, the state has annually raised the thresholds for each ranking.

**Example of Criteria 1.b.** (See pages A-2 - A-9.) Description of comprehensive system of reporting student performance in North Carolina.

**Example of Criteria 1.c.** (See pages A-10 - A-12.) Newspaper article describes raising the criteria used to place schools in Florida on the "critical" list.

**Example of Criteria 1.c.** (See pages A-13 - A-16.) Texas has created a plan to increase accountability requirements steadily from year to year. The excerpts from the Texas Accountability Manual describe the steps over the next four years.

## Accountability Reports

- 2. Reports include summaries of performance of children disaggregated by:**
- a. income level.**
  - b. race.**
  - c. primary language.**
  - d. performance by quartile.**
  - e. continuing students (students who have been at the school for a specified time).**
  - f. special education.**

**Example of Criteria 2.a,b,c.** (See pages A-17 - A-20.) The Texas Education Agency produces reports that present performance summaries by race, primary language, and income for individual schools, districts and the state as a whole.

**Example of Criteria 2.d.** (See page A-21.) The Chicago Public School Accountability Department places on the internet a report listing each school's performance for the last seven years. The report lists the percentage of children scoring in each quartile.

**Example of Criteria 2.f.** (See pages A-22 - A-25.) Newspaper article describing the effects that will occur from Texas, including performance of special education students in total score of school.

## Accountability Reports

### 3. Reports compare performance to:

- a. average performance of all students in state or district.
- b. average performance of students with comparable demographics.
- c. benchmark performance for students with comparable demographics.  
Benchmark scores come from the top 10%-15% of the schools in the same demographic group during the previous year.

**Rationale.** Benchmark levels are the levels achieved by the highest performing schools within each particular category, usually based on income level. School performance reports can be made particularly meaningful to the public when performance is benchmarked against the achievement of the top 10 - 15% of schools with comparable students as well as to the overall top 10 - 15% of schools.

The percentage of children in the school included in the testing is very important to consider, because when many children are not tested, the school ranking is questionable.

The percentage of children who are continuing students (have been at the school for the bulk of their education experience) is an important factor to be considered in judging the effectiveness of a school program, particularly urban schools with high mobility.

The gains made by students from one year to the next are particularly important in judging schools in which student performance is low.

### **Challenges to be met and situations to be avoided**

- Some states use scores that are not readily clear to parents, for example using a scoring range that is unfamiliar to parents (scores which range between 213 and 240). Parents do not readily understand the significance of these scores.
- Schools that have special populations (i.e., magnet schools that draw high performers from many neighborhoods) should not be compared to neighborhood schools, but rather to schools with comparable populations.

**Example of Criteria 3.c. (See pages A-26 - A-27.)** Description of instances in which benchmark levels of performance were established in schools serving high socio-economic populations.

**Example of Criteria 3.c. (See pages A-28 - A-33.)** Just For Kids, a non-profit organization, prepares and distributes data on benchmark performance in Texas Schools.

## Accountability Reports

### 4. Reports include:

- a. the percent of students in the school taking the assessments and included in reports.
- b. percent of students exempted for particular reasons, i.e., special education, language or other criteria.
- c. analysis of performance on individual topics and standards, if possible.

**Example of Criteria 4.a.** (See page A-34.) Virginia Standards of Learning requires reporting on the percent of students with disabilities and the percent of limited English proficient students exempted from testing.

**Example of Criteria 4.b.** (See pages A-35 - A-36.) Texas Education Code (Sec. 39.075) requires reports on the percent of students taking the test and the percent of students exempted. Reports are for individual schools, districts and the state as a whole.

**Example of Criteria 4.b.** (See pages A-37 - A-38.) Texas Education Agency is to investigate instances of excessive numbers of students being exempted from the testing.

**Example of Criteria 4.b.** (See page A-39.) Article from "*Catalyst*" magazine in Chicago describes effects of large scale retention on test scores.

## **B. Rewards and Sanctions**

### **B. Rewards and Sanctions**

**1. Criteria for rewards/recognition are based on:**

**OPTIONS:**

- a. a pass rate set by the state, or**
- b. performance of students in relation to benchmark levels for comparable students/for all students, or**
- c. rate of improvement in relation to benchmark rates of improvement (based on actual highest rates of improvement attained with populations of comparable demographics).**

**2. Examples of rewards/recognition:**

- a. financial rewards to schools/teachers/principals/district administrators based on student performance (e.g., gains or maintenance of benchmark levels).**
- b. release from most state requirements.**

**3. Criteria for sanctions (critical for low performing schools):**

**OPTIONS:**

- a. a failure rate set by the state, or**
- b. lack of acceptable gains in student achievement in relation to average rates of improvement attained with:**

**OPTIONS:**

  - population of comparable demographics, or**
  - all disaggregated groups combined.**
- c. imposed only after intensive help has been provided.**

**4. Examples of sanctions:**

**OPTIONS:**

- a. reconstitution**
- b. expedited dismissal/transfer of low performing staff.**
  - teachers**
  - principals**
  - district administrators**

## **Rewards**

**1. Criteria for rewards/recognition are based on:**

**OPTIONS:**

- a. a pass rate set by the state, or
- b. performance of students in relation to benchmark levels for comparable students/for all students, or
- c. rate of improvement in relation to benchmark rates of improvement (based on actual highest rates of improvement attained with populations of comparable demographics).

**2. Examples of rewards/recognition:**

- a. financial rewards to schools/teachers/principals/district administrators based on student performance (e.g., gains or maintenance of benchmark levels).
- b. release from most state requirements.

**Rationale.** Benchmark levels are the highest levels attained by the top 10 - 15% of the schools in a group, usually defined by income level. All schools not at benchmark levels for their group should be expected to improve. Improvement is expected until school reaches the benchmark levels of performance. Note that the benchmark level is not raised every year but would be stable for several years. The reason is that on an annual basis 100% of the schools cannot score at the level of the top 10 - 15% of the schools.

### **Challenges to be met and situations to be avoided**

- A number of school districts use the percent of children at grade level on a standardized test (50th percentile) to judge schools. Limiting the criteria for acceptable improvement to the percent of students scoring at grade level can result in schools focusing on only part of the students in the school, i.e., those students fairly close to grade level. This criteria could also be insensitive to schools dealing with at-risk students and with high turnover.
- Limiting testing to just three or four grade levels may result in schools not providing intensive enough instruction in the grades that are not tested.
- Establishing improvement levels that are too low can lead to a complacency with low performance. Several states established low levels to get off the state's critical school lists (lists of schools which have mandated improvement and threat of greater intervention).

- Improvement in low performing schools should be significant, but at the same time realistic. Large gains on standardized measures are possible very quickly in kindergarten and grade 1, but not in upper grades. In upper grades, other assessments can be used during a phase-in period in addition to state criteria (for example, scores on oral reading tests or tests sensitive to gains on basics such as the MAST test).

**Example of Criteria 2.a.** (See pages B-1 - B-2.) Texas Education Code Sec. 39.111. Requires that student performance must be a factor in awards to schools.

**Example of Criteria 2.a.** (See page B-3.) Texas Spotlight Schools program rewards 12 schools in state that produced very high reading performance with low income populations.

**Example of Criteria 2.a.** (See pages B-4 - B-11.) Articles from *Education Week* on teacher compensation and student performance.

**Example of Criteria 2.a.** (See pages B-12 - B-16.) Reports from Consortium for Policy Research on Education on teacher compensation and student performance.

## **Sanctions**

### **3. Criteria for sanctions (critical for low performing schools):**

#### **OPTIONS:**

- a. a failure rate set by the state, or**
- b. lack of acceptable gains in student achievement in relation to average rates of improvement attained with:**

#### **OPTIONS:**

- population of comparable demographics, or**
- all disaggregated groups combined.**
- c. imposed only after intensive help has been provided.**

### **4. Examples of sanctions:**

#### **OPTIONS:**

- a. reconstitution.**
- b. expedited dismissal/transfer of low performing staff:**
  - teachers.**
  - principals.**
  - district administrators.**

## **Challenges to be met and situations to be avoided**

- School districts are placing the responsibility for improving performance solely on school personnel. Experience of implementors of research-based models point out the lack of consistency that has characterized the implementation of these models. In some schools, gains are significant; while in other schools, gains are minimal or non-existent. The failure to produce gains can be directly tied to failure of the school to implement critical components of the model.
- District leaders should be accountable for creating a support and accountability system that is able to lead to schools uniformly replicating the gains that have been achieved with the research-based models in comparable schools.

**Example of Criteria 4. (Section 39.054) See Texas Education Code.** Requires a performance report that focuses on student achievement being the basis for school and district planning and for the evaluation of the district superintendent and school principals (and directors of regional education service centers).

**Example of Criteria 4. (Section 39.131) See Texas Education Code.** Accreditation Sanctions.

**Example of Criteria 4.a. (See pages B-17 - B-21.)** Article from *Catalyst* magazine in Chicago describing new agreements between Chicago Public Schools and union regarding reconstitution.

**Example of Criteria 4.a. (See pages B-22 - B-24.)** Article from *Education Week* about reconstitution and its effects in several cities.

**Example of Criteria 4.a. (See pages B-25 - B-27.)** Report on study of reconstitution conducted by Consortium for Policy Research at the University of Wisconsin.

## **C. Improvement Requirements and Assistance**

- 1. Improvement plans are to be implemented if student performance is significantly below:  
OPTIONS:
  - a. benchmark level of achievement for any demographic group, or**
  - b. average level of achievement for any demographic group, or**
  - c. average level of achievement for all demographic groups combined.****
- 2. Requirements for improvement:
  - a. are rigorous, but reasonable (based on actual rates of improvement attained by the top 10% of schools with comparable demographics).**
  - b. include expectations for improvement for all students (students in all quartiles and students who had been exempted from the test).****
- 3. Information and recommendations on instructional materials, instructional practices and service providers are readily available and based on reliable research and the experience of high performing educators.**
- 4. Funding and policies provide schools with financial means, flexibility and capability to adopt programs that have produced high levels of achievement with schools with comparable populations (e.g., schools can obtain support from any providers who can document effectiveness of their services).**
- 5. There is a system in place to evaluate the performance of “service providers.” This evaluation includes reports on increases in performance of students in schools that used the “service providers.”**

## **Improvement Requirements and Assistance**

**1. Improvement plans are to be implemented if student performance is significantly below:**

**OPTIONS:**

- a. benchmark level of achievement for any demographic group, or**
- b. average level of achievement for any demographic group, or**
- c. average level of achievement for all demographic groups combined.**

**Example of Criteria 1.** (See pages C-1 - C-2.) Excerpt from Texas Accountability manual on requirements for improvement.

**Example of Criteria 1.a.** (See pages C-3 - C-6.) Excerpt from 11/04/98 article "Strategies for Fixing Failing Schools," in Education Week commissioned by the PEW Foundation, includes discussion on what level of performance that schools should be held accountable to reach.

## **Improvement Requirements and Assistance**

### **2. Requirements for improvement:**

- a. are rigorous, but reasonable (based on actual rates of improvement attained by the top 10% of schools with comparable demographics).**
- b. include expectations for improvement for all students (students in all quartiles and students who had been exempted from the test).**

### **Challenges to be met and situations to be avoided**

- Although many states require intervention for the lowest schools, there are no requirements regarding interventions in relation to benchmark levels. Potential problem is that schools with low but not lowest achievement will not have a sense of urgency to improve.
- Most states begin state-level testing in 3rd or 4th grade. These tests are often not sensitive to gains made by students performing years below grade level. Assessments need to be obtained that are sensitive to gains made by low achieving students.

## Improvement Requirements and Assistance

**3. Information and recommendations on instructional materials, instructional practices and service providers are readily available and based on reliable research and the experience of high performing educators.**

**Rationale.** School personnel (principals and teachers) need reliable information on the materials and practices that are best suited for helping them educate the children at their school. Materials and practices that may be effective in one type of school setting (a school serving affluent children) may not be effective in another type of school setting (a school serving children from low income homes). School personnel need information on what works in their situation. **NOTE: Schools, districts and states need ready access to practical information based on trustworthy research.**

Reports on instructional materials and practices that have produced high levels of achievement should include clear statements of what is required for the results to be replicated and the specific types of students for which the instructional materials and practices have been found to be effective. Untested innovations are not recommended except for the purpose of evaluating them experimentally. Large scale implementations of an approach only occur when there is reliable data to indicate the effectiveness of the approach and there is sufficient capability (including funding for start-up costs) to replicate its results. Ongoing effectiveness data on instructional programs is aggregated across schools. Data on student achievement should be reported for each major instructional program and made widely available to educators and the public.

### Challenges to be met and situations to be avoided

- District and state criteria for the adoption of instructional materials are not based on proven effectiveness. Criteria established often are based on the latest education fad, rather than on proven and reliable research. This situation has resulted in programs that have proven to be effective with the at-risk students being excluded from lists.
- Refraining from recommending innovations that by definition have not been tested does not prohibit districts from experimenting with innovative practices. However, the decision to experiment should be made by the district. Encouragement by district offices, regional centers, and departments of education to innovate should be accompanied by provisions for evaluating the experiment.

- Do not assume that educators who have been successful in a school serving affluent populations can be helpful to schools serving high numbers of at-risk students.
- Avoid assigning educators who have not successfully led an improvement effort to a low performing school.

**Example of Criteria 3.** As part of the Texas Reading Initiative, Texas has prepared and distributed to every school a booklet entitled "Beginning Reading Instruction Components and Features of a Research Based Reading Program," available from TEA Publication Divisions, 1701 North Congress Avenue, Austin, TX 78701; Pub. #CU710501.

**Example of Criteria 3.** California Reading Initiative: "The California Reading Initiative is a series of laws based on reputable research summaries that should guide effective practice in reading instruction. The laws were passed with several intents: to produce higher achievement scores on a statewide basis; to use research-based practices to prevent reading failure; to change current practice in reading education so that it is more compatible with research-based guidelines; and to ensure consistency and continuity of instruction for children as they move through the grades or from school to school." For more information, write: Comprehensive Reading Leadership Center, Sacramento County Office of Education, 9738 Lincoln Village Drive, Sacramento, CA 95827; fax (916) 228-2404.

**Example of Criteria 3.** California Language Arts Framework:

<http://www.cde.ca.gov/cilbranch/eltdiv/rlafw.htm>

The California Language Arts Framework incorporates rigorous research findings into its recommendation for beginning reading instruction.

**Example of Criteria 3.** (See pages C-7 - C-8.) The American Institutes for Research is currently producing a report on prominent and promising comprehensive school reform approaches, *Judging Schoolwide Reform: A Teacher and Administrator Guide to 26 Noteworthy Approaches*. The report, which will be available in January of 1999 is commissioned by: the American Association of School Administrators, American Federation of Teachers, National Association of Elementary School Principals, National Association of Secondary School Principals, and National Education Association. In the report, the institute describes 26 reform approaches, critically evaluates the research bases, and summarizes implementation and costs associated with each program. The report is intended to guide schools and districts in making informed decisions regarding school reform strategies.

**Example of Criteria 3.** Virginia Advisory Board to provide assistance to schools is composed of educators from schools scoring at the top 10% of all schools in state and includes representatives from low income and urban and rural schools.

**Virginia Standards of Learning 20 131 310**

C. The Board of Education shall establish a Peer Educator Advisory Group to provide technical assistance in evaluating corrective action plans. The advisory group shall consist of fifteen (15) educators with five (5) The representatives each from urban, suburban, and rural schools. Representatives shall be academic classroom teachers and principals from elementary, middle, and secondary schools and shall be selected from among the top ten percent of schools in each category on the state SOL tests. Each member of the committee shall serve for no more than two years. Terms of service shall be designed to provide continuity to the group as a whole.

## **Improvement Requirements and Assistance**

**4. Funding and policies provide schools with financial means, flexibility and capability to adopt programs that have produced high levels of achievement with schools with comparable populations (e.g., schools can obtain support from any providers who can document effectiveness of their services).**

**Example of Criteria 4.** Virginia Standards of Learning require schools to provide extra time on reading for students not reading adequately.

### **Virginia Standards of Learning**

VAC 20-13-80 C.... Students who are not successfully progressing in early reading proficiency and/or who are unable to read the material necessary for instruction with comprehension shall receive additional instructional time in reading.

VAC 20-131-160. Additional Reading Instruction. Each school shall ensure that students who are unable to read with comprehension the materials necessary for instruction receive additional instruction in reading, which may include summer school.

**Example of Criteria 4.** (See pages C-9 - C-12.) The Consortium for Policy Research in Education (CPRE) at the University of Pennsylvania is studying the role of districts in supporting the implementation of reforms. Preliminary findings are presented of experiences in several cities.

**Examples of Criteria 4.** (See page C-13.) A report by the National Education Goals Panel in November 1998 "Exploring Achievement gains in North Carolina and Texas " included a description of increased authority and flexibility given by the state to districts and schools.

**Example of Criteria 4.** The Rodeo Foundation in Houston, Texas, has donated several million dollars to implement a research-based program from Wesley Elementary in 20 schools. Wesley serves a low income population, but has consistently produced high scores on the state test. Each school will receive the services of a full-time teacher for its first year or two in the project.

**Example of Criteria 4.** Several districts in California adopted a research-based beginning reading program. The Packard Foundation donated funds to each district to create a cadre of master teachers who support teachers as they learn to use the new curriculum.

**Example of Criteria 4.** (See pages C-14 - C-17.) Article commissioned by the PEW Foundation discusses role of district leader in improving schools.

## **Improvement Requirements and Assistance**

**5. There is a system in place to evaluate the performance of “service providers.” This evaluation includes reports on increases in performance of students in schools that used the “service providers.”**

### **Challenges to be met and situations to be avoided**

- Implementing instructional models in a manner that produces uniform success from school to school requires creation of a carefully coordinated support and accountability structure. Part of this effort revolves around ensuring that service providers offer effective services. Evaluation of service providers is rarely found.

**Examples of Criteria 5.** (See pages C-18 - C-26.) The "Catalyst" magazine produced a series of articles regarding satisfaction with external service providers in Chicago.

## **D. Interventions for Low Performing Schools**

**Note: These options are in addition to those described in Section C, “Improvement Requirements and Assistance.”**

- 1. Low performing schools are required to construct improvement plans that are inclusive, preventative, intensive, and begin in preschool and kindergarten.**  
**OPTIONS:**
  - **School constructs own construction plan.**
  - **School implements research-based model.**
  
- 2. Guidance is provided to school.**  
**OPTIONS:**
  - **by district-appointed consultant.**
  - **by consultant that school is able to choose.**
  - **by consultant who has produced high level of performance in similar school.**
  
- 3. Efforts are made to develop teacher involvement and support for the intervention plans:**
  - a. vote of staff.**
  - b. facilitate transfer of oppositional teachers.**
  - c. incentives to attract and retain effective teachers (teachers who have brought about reasonable levels of student achievement).**
  
- 4. On-going assessment plan is constructed to determine if students are making reasonable progress. Assessments are administered periodically through the school year. Periodic reviews by district administrator determines the extent to which program is being implemented as planned; if components of the improvement plan are not being implemented effectively, the administrator takes the necessary actions.**
  
- 5. When a school’s improvement plan does not produce satisfactory achievement gains after two years, stronger interventions are required:**  
**OPTIONS:**
  - a. Transfer/dismissal of staff members.**
  - b. Reconstitution. (See Sanctions.)**

## **Interventions for Low Performing Schools**

- 1. Low performing schools are required to construct improvement plans that are inclusive, preventative, intensive, and begin in preschool and kindergarten.**

**OPTIONS:**

- a. School constructs own construction plan.**
- b. School implements research-based model.**

**Example of Criteria 1.** (See pages D-1 - D-2.) The American Federation of Teachers has published a document entitled "Raising Student Achievement—A Resource Guide for Redesigning Low Performing Schools." The document describes models that it considers to be promising. The document describes models that it considers to be promising. Available from the American Federation of Teachers, 555 New Jersey Avenue, N.W., Washington, DC 20001; ask for item 370, 7/97. Website: [www.aft.org](http://www.aft.org)

## **Interventions for Low Performing Schools**

**2. Guidance is provided to school.**

**OPTIONS:**

- a. by district-appointed consultant.**
- b. by consultant that school is able to choose.**
- c. by consultant who has produced high level of performance in similar school.**

### **Challenges to be met and situations to be avoided**

- Do not assume that educators who have been successful in a school serving affluent populations can be helpful to schools serving high numbers of at-risk students.
- Avoid assigning educators who have not successfully led an improvement effort to a low performing school.

**Example of Criteria 2.** (See pages D-3 - D-4.) Virginia uses educators from high performing schools to assist schools with their improvement plans.

**Example of Criteria 2.** (See pages D-5 - D-13.) Article in *Education Week* discusses efforts of several cities and states to provide guidance to low performing schools.

## **Interventions for Low Performing Schools**

- 3. Efforts are made to develop teacher involvement and support for the intervention plans:**
- a. vote of staff.**
  - b. facilitate transfer of oppositional teachers.**
  - c. incentives to attract and retain effective teachers (teachers who have brought about reasonable levels of student achievement).**

**Example of Criteria 3.** (See pages D-14 - D-15.) North Carolina Department of Education Press Release describing incentives for teachers in low performing schools that show improvement.

**Example of Criteria 3.** (See pages D-16 - D-19.) Newspaper article describing potential disincentives of reform efforts in retaining or attracting good teachers in low performing schools.

## **Interventions for Low Performing Schools**

- 4. On-going assessment plan is constructed to determine if students are making reasonable progress. Assessments are administered periodically through the school year. Periodic reviews by district administrator determines the extent to which program is being implemented as planned; if components of the improvement plan are not being implemented effectively, the administrator takes the necessary actions.**

**Example of Criteria 4.** (See pages D-20 - D-21.) Several instructional models have replicated effectiveness in improving reading performance in low performing schools. The improvement is dependent on the level of implementation. A summary of the components of one model are listed as well as a graph showing the relationship between lessons taught and student performance on tests.

## **Interventions for Low Performing Schools**

**5. When a school's improvement plan does not produce satisfactory achievement gains after two years, stronger interventions are required:**

**OPTIONS:**

- a. Transfer/dismissal of staff members.**
- b. Reconstitution. (See Sanctions.)**

**Example of Criteria 5.** (See page D-22.) Article commissioned by the PEW Foundation discusses role of reconstitution.

## E. Criteria for Assessment

### 1. Assessment quality:

- a. aligned with standards, including extended coverage on most important content.
- b. valid.
- c. reliable.
- d. sensitive to students' gains made during school year, including students who are below grade level.
- e. provides clear measure of domain in elementary grades by avoiding performance from other domains (e.g., the reading test is not turned into a writing test).

### 2. Assessment coverage:

- a. begins in lower grades.
- b. administered in each grade level in key subjects.
- c. inclusive, with maximum numbers of students taking test.

### 3. Assessment scores:

- a. passing level of assessments are rigorous but reasonable: indicating readiness for next level (e.g., determine correlation between passing 8th grade math test and passing algebra).
- b. a range of levels is utilized to describe student performance (e.g., proficient, satisfactory, basic, below basic).
- c. criteria for student score levels:  
    **OPTIONS:**
  - set high when assessment is established, or
  - set lower at beginning and then raised systematically over time.

### 4. Assessment administration:

- a. designed and scheduled so that they do not take excessive time from instruction.
- b. results returned prior to the beginning of the next school year.
- c. tests administered late in the school year to make expectations for accountability clear.

## **Criteria for Assessment**

### **1. Assessment quality :**

- a. aligned with standards, including extended coverage on most important content.**
- b. valid.**
- c. reliable.**
- d. sensitive to students' gains made during school year, including students who are below grade level.**
- e. provides clear measure of domain in elementary grades by avoiding performance from other domains (e.g., the reading test is not turned into a writing test).**

**Rationale.** The assessment is aligned with the content standards. There is an explicit relationship between those content standards that require extensive time to each and/or are most important and what is covered in the assessment. For example, symmetry and fractions both appear in elementary grades. Fractions though are a critical component of more advanced content and should receive greater emphasis on the assessment. Basics in reading involve reading with accuracy and fluency and answering literal and inferential comprehension question. In math, basics involve the student doing operations and applying them to problem solving situations that can be translated to simple equations.

### **Challenges to be met and situations to be avoided**

- The absence of priorities in most content standards places the issue of what is important in hands of test preparers. If priorities do not appear in standards, careful monitoring should occur over the test preparation process in order to ensure that the test reflects appropriate coverage of knowledge and skills.
- Tests made up solely of challenging applications of the standards may not be sensitive to performance gains made by at-risk students in learning the content, skills and direct applications of the standards. In some states, the percentage of students passing state tests is very low, because the tests include only the advanced applications.
- Assessments on which students must apply knowledge from other areas undermines the assessment of targeted content standards. For example, a report from the CRESST center in Los Angeles (<http://www.cse.ucla.edu>) pointed out that at-risk students' scores on math assessments may not be reflective of their

true math ability because of relatively complex language requisites of the test. Another example is reading tests which require a great deal of writing.

- If performance assessments are used in areas other than writing, they must be (a) scored quantitatively and reliably, (b) cost effective in terms of time devoted to the test and resources to grade tests, and (c) reflective of the scope of content standards. (RESOURCES: Sometimes performance assessments are so expensive to score reliably that the overall assessment system becomes too expensive. TIME: Sometimes performance assessments take so much time for students to complete, there is not time for conventional tests to sample the range of core content. SCOPE OF CONTENT: The performance assessment items on some tests result in teachers narrowing the scope of the content they teach through the school year, not giving adequate attention to the broad core content. This also relates to aligning the assessment with all the priorities of the content standards, not just one or two. For example, if a U.S. history performance test deals only with the Civil War, teachers might spend a disproportionate amount of time on the Civil War.)
- At-risk students in the late elementary grades and beyond can be so far below grade level that the assessment will not be sensitive to gains they might be making, which discourages teachers and students. The challenge is to create options that might allow students who score very poorly to take the test for a lower grade level to establish their actual performance level and to measure any gains that might have occurred during the school year. An analogous procedure could be enacted for students achieving far above grade level.

**Example of Criteria 1.a. (See pages E-1 - E-9.)** Virginia produced blueprints for state test that clearly communicate how assessments are aligned with standards. Copies are available from the Commonwealth of Virginia Department of Education, James Monroe Building, 101 N. 14th Street, Richmond, VA 23129. Phone: (804) 225-2102.

## Criteria for Assessment

### 2. Assessment coverage:

- a. begins in lower grades.
- b. administered in each grade level in key subjects.
- c. inclusive, with maximum numbers of students taking test.

Testing of reading and math begins in the primary grades and occurs in each grade throughout elementary grades. Other subject content (social studies, science, writing) can be tested, beginning at third or fourth grade. Testing in grade one and two can be diagnostic and not necessarily part of the accountability system.

### Challenges to be met and situations to be avoided

- Reading performance of children in first grade correlates with reading achievement in later grades. The need for assessment of student progress in first grade is very important.

**Example of Criteria.** (See pages D-10- D-13.) Description of testing program in North Carolina.

**Example of Criteria 1.a.** Virginia and Texas have developed diagnostic tests to be used in kindergarten through second grade. Texas Primary Reading Inventory is available from the Texas Education Agency. In Virginia, the state department has produced tests, PALS-I and PALS-II (Phonological Awareness and Literacy Screening Test ) which are available from the State Department of Education in Richmond, Virginia.

**Examples of Criteria 1.c.** Two articles from the Council for Exceptional Children discuss issues on Inclusive Assessments and Accountability Systems.

**Examples of Criteria 1.c.** (See pages D-14- D-23.) Two articles from the Council for Exceptional Children on Inclusive Assessments and Accountability Systems.

## Criteria for Assessment

### 3. Assessment scores:

- a. passing level of assessments are rigorous but reasonable: indicating readiness for next level (e.g., determine correlation between passing 8th grade math test and passing algebra).
- b. a range of levels is utilized to describe student performance (e.g., proficient, satisfactory, basic, below basic).
- c. criteria for student score levels:  
OPTIONS:
  - set high when assessment is established, or
  - set lower at beginning and then raised systematically over time.

**Rationale.** Ideally, the process for establishing performance levels (e.g., basic, proficient, advanced) is based on actual student performance rather than solely on the judgment of an advisory committee. Passing levels should correlate with performance at subsequent levels as indicated by scores on subsequent course work, scores on state assessments in later grades, the need for remedial course work in later grades, and scores on national norm-referenced test(s).

### Challenges to be met and situations to be avoided

- Passing levels that are too low: In one state, passing the state's eight grade math test did not correlate with passing algebra. However, reaching the mastery level on the state's eighth grade math test did correlate with passing algebra. This example suggests that the passing level is too low, but the mastery level is not.
- In one state, the intended results for writing include spelling and grammar, yet students still can receive an excellent rating on the state assessment even though their papers are full of spelling and grammar mistakes. Most communities would prefer that grammar and spelling be counted as part of the score or that two scores be given, one for content and one for spelling and grammar.

**Example of Criteria.** (See pages D-24 - D-25.) Newspaper article describes reaction in Florida to new test that is more difficult.

**Example of Criteria.** The federal NEAP includes four levels to describe student performance.

**Example of Criteria.** Texas legislation requires comparison to national performance. Texas Education Code, Sec. 39.028. Comparison of state results to national results.

The state assessment program shall obtain nationally comparative results for the subject areas and grade levels for which criterion-referenced assessment instruments are adopted under Section 39.023.

(g) The State Board of Education may adopt one appropriate, nationally recognized, norm-referenced assessment instrument in reading and mathematics to be administered to a selected sample of students in the spring.

If adopted, a norm-referenced assessment instrument must be a secured test. The state may pay the costs of purchasing and scoring the adopted assessment instrument and of distributing the results of the adopted instrument to the school districts. A district that administers the norm-referenced test adopted under this subsection shall report the results to the agency in a manner prescribed by the commissioner.

## **Criteria for Assessment**

### **4. Assessment administration:**

- a. designed and scheduled so that they do not take excessive time from instruction.**
- b. results returned prior to the beginning of the next school year.**
- c. tests administered late in the school year to make expectations for accountability clear.**

### **Challenges to be met and situations to be avoided**

- Testing requirements can interfere with instruction when schools have to administer two different types of tests during the school year. If the state test and another test that is dissimilar are administered and the tests are high stakes, a good deal of instructional time may be taken from the regular program for test preparation.
- One state had a format that was very different from the tasks students typically did in school. The students were given multiple test items in which one or more of the items were correct answers.
- Assessments administered too early in the spring of the school year will not give teachers adequate time to address the content standards for that grade level.

**Examples of Criteria.** Report on North Carolina and Texas Accountability systems describes computerized feedback system.

Scores on the tests are provided to students, parents, teachers, schools and school districts. Both states have a well-designed computerized system of storing the testing information and providing access to it in various ways to teachers, principals and school districts. Tests are graded centrally in Texas and in regional locations in North Carolina. Access to school level results is provided on the Internet in both states. Both states have developed varied formats for reporting test results by sets of questions related to key learning objectives. Teachers in both states have access to summaries and individual tests of students entering their classes each year. In both states test items are made available after each test to the public to counter criticism of bias in test items. (p. 22)

## **F. Content Standards**

- 1. Rigorous yet reachable.**
- 2. Measurable, clearly and simply specified at each grade level.**
- 3. Includes important content with the most important subject matter identified:**
  - OPTIONS:**
  - a. in the standards, or**
  - b. in the assessment blueprint.**
- 4. Neither dictate pedagogy nor contradict findings of rigorous research.**
- 5. Balance of knowledge and skills with the basics specified for the primary grades.**

## Content Standards

### Criteria 1. Rigorous yet reachable

Standards are appropriately challenging, important and commensurate with standards of other states and nations. The standards include a balance and link among skills, knowledge and application. The standards are reasonable in terms of difficult level and in terms of quantity of information.

Three organizations have evaluated the rigor of state standards and assigned ratings. These ratings and the criteria used by the organizations can be found in these documents:

- Great Expectations—Defining and Assessing Rigor in State Standards for Mathematics and English Language Arts. Council of Basic Education, January, 1998. E-mail: [info@c-b-e.org](mailto:info@c-b-e.org)
- Defining World Class Standards—American Federation of Teachers. <http://www.aft.org>
- The State of State Standards—Thomas Fordham Foundation, July 1988. <http://www.edexcellence.net>

The ratings of these organizations can be confusing since the criteria of the organizations differ.

### Challenges to be met and situations to be avoided

- At the present time, standards reflect the best estimate by adults with a range of backgrounds of what students should be able to do. Over time standards should be revised to reflect measures of what is rigorous yet reachable. For example determining what is reasonable for students to know at a particular grade level could be determined by observing what in fact students are able to achieve in high achieving schools. If a high percentage of students in the higher performing schools are able to do the tasks called for by the standards, the standards would be considered reasonable for that grade level. If the students can do more than what is specified in the standards, the standards can be raised.
- U.S. students are very diverse. High student performance standards for a school serving an affluent neighborhood would not be appropriate for a school serving children of poverty.

## **Example of Criteria**

California Math Standards, 1997

Web Site: [http://www.cde.ca.gov/board/k12math\\_standards.html](http://www.cde.ca.gov/board/k12math_standards.html)

The Math Standards passed by the California State Board in 1997 represent an attempt to make expectations for U.S. students commensurate with those of countries that have scored at higher levels on international tests. The new standards are designed to prepare students for algebra in eighth grade, followed by geometry in ninth grade and advanced algebra in tenth grade. The expectations in grades 1-7 have been increased in order to prepare students for the more challenging standards of grade 8 and above.

## **Example of Criteria (See pages F-1 - F-2.)**

*Education Week* special edition "Quality Counts" included a summary of the review of state standards using information provided by the Council for Basic Education and the American Federation of Teachers.

## **Content Standards**

### **2. Measurable, clearly and simply specified at each grade level.**

Standards are written to clearly communicate expectations to educators, publishers, families, and communities. Numerous examples are provided to exemplify the intent of the standard. Standards are written clearly enough that preparers of assessments have clear guidelines as to what content is to be tested.

The advantage of including standards for every grade level is that everyone—administrators, teachers, students, parents, and other interested citizens—knows what is expected of students during each school year. Setting standards for every grade level simplifies the process of aligning curriculum to standards. It also eliminates confusion among teachers about what they can expect students to have learned the year before and what they must prepare their students to be ready to learn the following year. Parents will know exactly what material their children are learning and can participate readily in their children's education. Students also will be privy to a clear set of expectations for each academic year

### **Examples of Criteria**

The states of Virginia and California have produced content standards that are simple and clearly stated. The standards have received high rankings from several sources.

## **Content Standards**

**3. Includes important content with the most important subject matter identified:**

**OPTIONS:**

- a. in standards, or**
- b. in the assessment blueprint.**

Priorities can communicate the relative importance of each standard and the amount of effort that will be required to teach students to reach the standard. Priorities are particularly important in the primary grades.

### **Challenges to be met and situations to be avoided**

- The absence of priorities in content standards places the issue of what is important in hands of test preparers. If priorities do not appear in standards, careful monitoring should occur over the test preparation process in order to ensure that the test reflects adequate coverage of priority knowledge and skills. The test specifications with the number of items for each standard indicates the priorities and should be widely distributed before the final assessment is completed.

### **Example of Criteria**

The Virginia Standards of Learning are not prioritized; however, the state has distributed blueprints for state tests. The blueprint, by specifying the number and type of problems that will be on the state test, is in fact setting priorities. Copies are available from the Commonwealth of Virginia Department of Education, James Monroe Building, 101 N. 14th Street, Richmond, VA 23129. Phone: (804) 225-2102. [http:](http://)

## **Content Standards**

**4. Neither dictate pedagogy nor contradict findings of rigorous research.**

### **Challenges to be met and situations to be avoided**

- Content Standards tell what students are to be able to do, not how teachers are to teach.
- Content Standards are not written so that the sequence in which teachers present topics is consistent with research findings on teaching and learning.

### **Example of Criteria**

The California Language Arts Standards incorporate the rigorous research on beginning reading. Available from: Publication Division, Sales Office, California Department of Education, P.O. Box 271, Sacramento, CA 98512-0271; Fax 916-323-0823.

These research findings and their implications were summarized in a document on beginning reading, "Every Child Reading," published by the Learning First Alliance, a coalition of major teacher organizations, administrator organizations, and other education organizations. <http://www.learningfirst.org>

## Content Standards

5. **Balance of knowledge and skills with the basics specified for the primary grades.**

Content standards should specify the core skills and content that need to be taught in kindergarten and the primary grades. This specification is important in communicating the expectation that teaching is to be organized so that all children do learn the basics in the primary grades. The desire of educators to ensure that children are able to think analytically and communicate well has sometimes resulted in teachers not placing enough emphasis on core content such as decoding skills, math facts and basic sentence and paragraph construction mechanics.

### Example of Criteria

**California Mathematics Standards.** The California Mathematics Standards clearly specify that critical numeration skills (e.g., facts, reading and writing numbers, place value) be taught in early grades.

## **G. Responsibilities of Other Organizations**

### **Legislators, School Boards and Teacher, Parent, Community, and Business Groups**

**These organizations establish and maintain an active quality assurance role in the process of adopting and implementing standards-based reform.**

### **Federal Government**

**The federal government provides funds for high quality research on important issues and funds efforts to support research-based improvement efforts for at-risk groups—students with disabilities, children of poverty, and English learners.**

## **Legislators, School Boards and Teacher, Parent, Community, and Business Groups**

**These organizations establish and maintain an active quality assurance role in the process of adopting and implementing standards-based reform.**

The organizations authorize and fund external independent audits to determine the extent to which assessment efforts are valid and reliable and the extent to which improvement efforts are research-based and well implemented. School boards require reports from administrators regarding programs selected for improvement efforts and periodic reports on students' progress and performance. Stakeholders include: elected officials, the public, business groups, professional groups for teachers and administrators, and school boards.

The Texas Accountability Story describes an excellent example of how business can be provide leadership in standards-based reform.

Texas—School Information Project—Just for the Kids Foundation compares schools against benchmark level performance.

Texas Association of School Boards publishes a document entitled "How Board Policy Can Support Research-Based Practices" and a document entitled "What Can Trustees do to Ensure Every Child Reads" that focuses on appropriately implementing research-based practices for teaching reading.

American Federation of Teachers has produced a document entitled "What Works" that lists school-wide models that have produced high levels of achievement in schools with high number of students at-risk.

A consortium of education organizations is exploring a process for validating research on educational approaches.

In an article by David Glovin and Jean Rimbach (1998) entitled "Court-ordered school reforms gain varied levels of acceptance," it was stated:

In its final ruling in a 28-year legal fight, the Supreme Court in May told New Jersey to intensify efforts to improve education in the state's 28 poorest districts, including Paterson, Passaic, and Garfield. The justices demanded that the state provide schools with adequate money and require changes in instructional methods.

The ruling was hailed as a landmark by many urban education advocates, who see the decision as a way to bring state-of-the-art teaching techniques to some of New Jersey's most troubled schools.

The court ordered 50 elementary schools to undergo top-to-bottom reform by September, 100 more by September 1999, and 150 more the next year. The process, called whole-school reform, requires teachers to change the way they teach, the books they use, and the way they run classes.

The state wants schools to adopt the Success for All method. The program requires elementary children to spend 90 minutes a day reading in groups of 15, with frequent tests and ample tutoring. Schools, though, may use other methods, including one that has students doing work above grade level.

Assistant Education Commissioner Barbara Anderson says that at least 50 schools will be using one reform or another, although she concedes that many of those schools have already been using reform programs. "You'll have continuing schools and others that are new schools," she said.

## **Federal Government**

**The federal government provides funds for high quality research on important issues and funds efforts to support research-based improvement efforts for at-risk groups—students with disabilities, children of poverty, and English learners.**

- direct research, dissemination, and technical assistance funding toward important goals that educators and the public strongly support.
- improve the quality of research through its funding criteria, by systematically building on syntheses of existing research, and by studying highly effective programs and schools.
- require that funds be used for research-based practices in teacher preparation and professional development, in dissemination, and in technical assistance for teacher, schools, states, and parents.
- encourage districts to use Title I, IDEA, and bilingual funds to replicate proven instructional models.

## **G. Responsibilities of Other Organizations**

### **Legislators, School Boards and Teacher, Parent, Community, and Business Groups**

**Stakeholders establish and maintain an active quality assurance role in the process of adopting and implementing standards-based reform.**

### **Federal Government**

**The federal government provides funds for high quality research on important issues and funds efforts to support research-based improvement efforts for at-risk groups—students with disabilities, children of poverty, and English learners.**

## References

- Joftus, S. & Berman, I. (1998, Jan.). *Great Expectations?—Defining and assessing rigor in state standards for mathematics and English language arts.* (pp. 19-29). Council for Basic Education.
- Glaser, R. & Silver, E. (1994, June). *Assessment, testing, and instruction: Retrospect and prospect.* CSE Technical Report 379, p. 22. CRESST/Learning Research and Development Center, University of Pittsburgh.
- Grissmer, D., & Flanagan, A. (1998, Nov.). *Exploring rapid achievement gains in North Carolina and Texas.* (pp. 23-24). National Education Goals Panel.