



The Commonwealth Accountability Testing System

Research Report No. 312

Prepared by:

Lynn Aubrey, Kara Daniel, Erin McNees, Stacie Otto,
and Greg Hager, Ph.D., Committee Staff Administrator

The Commonwealth Accountability Testing System

Program Review and Investigations Committee

Program Review and Investigations Committee Staff Report

Greg Hager, Ph.D.
Committee Staff Administrator

PROJECT STAFF

**Lynn Aubrey
Kara Daniel
Erin McNees
Stacie Otto**

Research Report No. 312

Legislative Research Commission

Frankfort, Kentucky

Program Review and Investigations Committee

Adopted August 22, 2003

This report has been prepared by the Legislative Research Commission and printed with state funds.

Foreword

Program Review staff would like to acknowledge and thank Scott Trimble, Kevin Hill, Bill Insko, Cindy Owen, Julie Whitaker, Cherry Boyles, and Bonnie Brinly and other staff from the Kentucky Department of Education. Staff would also like to thank Audrey Carr, Sandy Deaton, Dudley Cotton, Dan Jacovitch, Rebecca Brickey, the Legislative Research Commission Project Center and other LRC staff who provided assistance with this study. Finally, staff would like to thank the members of the National Technical Advisory Panel on Accountability and Assessment, the Vermont Department of Education, and the teachers, principals, and superintendents who responded to our surveys.

Robert Sherman
Director

Frankfort, Kentucky
August 22, 2003

SENATE MEMBERS

David L. Williams
President, LRC Co-Chair

Richard L. Roeding
President Pro Tem

Dan Kelly
Majority Floor Leader

Ed Worley
Minority Floor Leader

Charlie Borders
Majority Caucus Chairman

Johnny Ray Turner
Minority Caucus Chairman

Elizabeth Tori
Majority Whip
Bob Jackson
Minority Whip



LEGISLATIVE RESEARCH COMMISSION

State Capitol

700 Capital Avenue

Frankfort KY 40601

502/564-8100

Capitol FAX 502-223-5094
Annex FAX 502-564-6543

www.lrc.state.ky.us/home.htm

Robert Sherman
Director

HOUSE MEMBERS

Jody Richards
Speaker, LRC Co-Chair

Larry Clark
Speaker Pro Tem

Gregory D. Stumbo
Majority Floor Leader

Jeffrey Hoover
Minority Floor Leader

Jim Callahan
Majority Caucus
Chairman

Bob DeWeese
Minority Caucus
Chairman

Joe Barrows
Majority Whip

Ken Upchurch
Minority Whip

MEMORANDUM

TO: The Honorable Paul E. Patton, Governor,
the Legislative Research Commission, and
Interested Individuals

FROM: Senator Katie Stine, Co-Chair
Representative Charlie Hoffman, Co-Chair

SUBJECT: Adopted Committee Staff Report: *The Commonwealth Accountability Testing
System*

DATE: August 22, 2003

In July 2002, the Program Review and Investigations Committee authorized a study of the Commonwealth Accountability Testing System (CATS). The committee approved the resulting report and adopted its recommendations on August 22, 2003. The Kentucky Department of Education's (KDE) response to the report is included as Appendix H.

In conducting the study, Program Review staff analyzed data and documentation KDE uses to measure educational progress under CATS. Interviews were conducted with staff of the department, the Kentucky Education Professional Standards Board, the Office of Education Accountability, and the Interim Joint Committee on Education. Staff attended a regional training session for scorers of writing portfolios and meetings of the National Technical Panel on Assessment and Accountability. Staff surveyed teachers, principals, and superintendents on issues related to CATS assessment and accountability.

The major conclusions of the report are that:

- KDE accepts dropout information from school districts without sufficient validation, so the accuracy of school, district, and statewide dropout rates is unknown.
- Program Review staff estimated that the CATS assessment cost school districts about \$10.6 million for the 2001-2002 school year. State costs were estimated at \$10.3 million for fiscal year 2003.
- Correlation between individual students' CATS and ACT scores is relatively high.
- Just over 50 percent of principals and superintendents responded to the Program Review staff's survey that the quality of education is better under CATS compared to the previous system. About a third of teachers said that education is better under CATS.
- There is little evidence from the survey to indicate that teachers are "teaching the test" if that is defined as teachers incorporating unreleased past questions into their lessons.
- More than half the teachers disagreed that the process used to set school improvement goals by the Kentucky Board of Education was appropriate.
- At least 75 percent each of teachers, principals, and superintendents responded to the survey that the accountability weights for reading, math, science, and social studies are about right. Around one-half of teachers responded that the weight for the writing portfolio was too high; about 30 percent of principals and superintendents agreed.
- Almost one-half of teachers responded to the survey that some teachers have biases that affect scoring of writing portfolios.
- Agreement rates between original scores assigned to writing portfolios and the scores assigned during audits of the writing portfolios do not appear to have increased consistently over recent years. During the writing portfolio audit, original scores assigned at the school level are revised downward more frequently than upward, indicating that original writing portfolio scores often may be inflated.

Based on these conclusions, the report recommended that the Kentucky Department of Education should:

- perform specific procedures to verify school districts' reported dropout statistics;
- consider sanctioning schools that underreport dropout statistics;
- implement a uniform student information system at the state level to track students;
- create a CATS testing expenditure category in the MUNIS system (the financial software package used for Kentucky school districts) and encourage school districts to utilize this category for all CATS administrative expenses;
- work with schools and districts to reduce, as much as possible, the practice of teachers scoring their own students' portfolios; survey teachers to determine how their portfolio scoring training can be improved; regularly replace benchmark portfolios with new samples; encourage schools to provide teachers with more opportunities to practice scoring writing portfolios;
- consider implementing a system to track the performance of portfolio scorers; and
- consider establishing consequences for schools that have low portfolio audit agreement rates.

Questions or requests for additional information should be directed to Greg Hager, Committee Staff Administrator for the Program Review and Investigations Committee.

Contents

Executive Summary	xi
Chapter 1: An Overview of the Commonwealth Accountability Testing System and This Study.....	1
Introduction	1
Description of This Study.....	2
How This Study Was Conducted.....	2
Organization of the Report.....	3
Major Conclusions	4
The Kentucky Education Reform Act	5
Goals of the Kentucky Education Reform Act	5
The Commonwealth Accountability Testing System.....	7
The CATS Test.....	8
Norm-referenced and Criterion-referenced Tests	9
Subjects Tested Under CATS	10
Types of Test Items.....	11
Multiple Choice	11
Open Response.....	11
Portfolios.....	12
Chapter 2: Kentucky’s Education Accountability System.....	13
Schools’ Two-year Goals	13
Calculation of Goals	13
Rewards and Assistance.....	15
Kentucky’s School Accountability Index.....	16
CATS Accountability and School Rewards	17
“Needs Assistance” Classification System.....	18
Dropout Rates and CATS.....	19
KDE Reports That the Dropout Rate Has Declined	20
Reporting Procedures for Dropouts	20
Self-reported Dropout Data Pose Potential Problems	21
<i>Recommendation 2.1</i>	23
<i>Recommendation 2.2</i>	23
The Importance of a Statewide Student Information System.....	23
<i>Recommendation 2.3</i>	24
Graduation Rates	25
Graduation Formulas Yield Different Results	26
The Financial Cost of the CATS Assessment	26
Local School Districts’ Costs.....	27
<i>Recommendation 2.4</i>	30

State Costs.....	30
Other States’ Assessment and Accountability Systems	31
State Testing.....	32
Types of Tests	32
Types of Questions Used in State Tests.....	32
Subjects Tested.	33
Accountability Systems	33
 Chapter 3: Evaluating CATS	35
The Surveys of Educators.....	35
The Effects of CATS on the Quality of Education	36
Comparison With the Previous Accountability System.....	36
Effect on Curriculum	37
Test’s Effect on Teaching	37
Effect on Student Learning	38
Core Content for Assessment.....	38
Repetition of Test Questions.....	38
KDE on Why the Repetition of CATS Questions Is Necessary	39
Setting Improvement Goals	40
Schools That Do Not Meet Improvement Goals.....	42
Consequences.....	42
Quality of Assistance	43
The Accountability Index	43
Student Accountability	45
Comparing CATS and ACT Scores	47
Trends in Statewide ACT and CATS Scores.....	48
The Correlation Between Students’ ACT and CATS Scores	49
KCCT and CATS Are Valid and Reliable According to NTAPAA	51
 Chapter 4: Writing Portfolios and On-demand Writing	53
Background	53
Kentucky Educators’ Opinions of the Writing Portfolio.....	56
A Comparison of Writing Portfolios and On-demand Writing	56
Training for Scoring of Writing Portfolios.....	57
<i>Recommendation 4.1</i>	59
Scoring Options	60
Monitoring of Portfolio Development and Scoring.....	62
Selection of Schools for Audit of Writing Portfolios	62
Auditing of Writing Portfolios.....	64
The Validity of the CTB Audit	66
Audit Results.....	66
<i>Recommendation 4.2</i>	69
<i>Recommendation 4.3</i>	69

Chapter 5: Kentucky’s Compliance With the No Child Left Behind Act	71
Remaining NCLB Compliance Issues	72
Accountability System	72
Baseline Calculation	73
Proficiency of All Students	74
Reducing Achievement Gaps	74
Only 16 States Hold Schools Accountable for	
Subgroup Performance	75
Assessments in Additional Subjects	75
Measuring Success	76
Concurrent Accountability Systems May Cause Confusion	76
Challenging Standards	76
Math and Reading Standards Need To Be Developed	
in Half the States	77
States Lower the “Proficiency” Bar	78
Academic Assessment	78
Thirty States Need To Add Assessments in Reading and Math	79
Inclusion of Students in the Assessments	79
Graduation Rates	81
Failing Schools and Districts	81
School Improvement	82
Corrective Action	82
School Restructuring	83
Only Eight States Have Timely Identification of Schools	
in Need of Improvement	84
NCLB Safe Schools	84
More Than 30 States Are in Compliance With NCLB Safe	
School Provisions	84
Teacher and Paraprofessional Qualifications	85
Kentucky Educators’ Opinions on Teacher Qualifications	86
Wisconsin Is the Only State With a Highly Qualified	
Teacher in Every Classroom	87
The Estimated Cost of Implementing NCLB	88
Works Cited	89
Appendix A: Summary of Statutes and Regulations Governing CATS	91
Appendix B: Kentucky Core Content Test, Sample Questions	113
Appendix C: “Needs Assistance” Schools	119
Appendix D: The Surveys of Teachers, Principals, and Superintendents	123
Appendix E: Calculating Correlations Between ACT and CATS Scores	151
Appendix F: Statement of NTAPAA on the Validity and Reliability	
of CATS	153
Appendix G: Provisions of the No Child Left Behind Act	159
Appendix H: Kentucky Department of Education’s Response to This Report	171

List of Tables

1.1	Timeline of Events Prior to KERA.....	6
1.2	Subjects Tested Under CATS	10
2.1	Percentage Weights of Components in the School Accountability Index by Grade Level.....	17
2.2	Estimate of School Districts' CATS Assessment Costs, 2001-2002	29
2.3	Estimate of State General Fund Expenditures for CATS Assessment, FY 2003 ..	31
2.4	Types of Test Items Used in States' Assessment Systems	32
2.5	State Report Cards That Report Student Performance Data by Groups	33
2.6	Information Used by States To Evaluate Schools.....	34
3.1	Educators' Opinions of the Weighting of Accountability Index Items	44
3.2	Kentucky Statewide Average ACT and CATS Scores for Public School Students Taking the ACT, 1999 to 2002	48
3.3	Correlation Between CATS and ACT Scores: Composite, Math, and Reading	50
3.4	Student-level CATS Scores Grouped by ACT Score	50
4.1	Writing Portfolio and On-demand Writing Scores, 1999 to 2002	57
4.2	Percentage of Time Teachers Report That Initial Writing Portfolio Scores Do Not Agree	61
4.3	Agreement Rates Between Original Scores and Scores Assigned During Audits for Writing Portfolios.....	67
4.4	Percent of Scores Changed in Writing Portfolio Audit by Scoring Method, 1998-1999 to 2001-2002	68
5.1	Where States Stand: NCLB Standards.....	78
5.2	Where States Stand: NCLB Teacher Qualification Provisions.....	87

List of Figures

1.A	How School Goals Relate to the Kentucky Core Content Test	7
2.A	A School's Long-term Growth Chart	14
2.B	2002 Classification of Schools.....	19
2.C	State Dropout Requirement and State Dropout Rate, 1993-2002	20
2.D	Statewide High School Graduation Rate Based on the NCES Formula, 1997-2002.....	25
3.A	Compared to the Previous System, the Quality of Education Is Now Better, About the Same, or Worse?	36
3.B	How Does the CATS Test Affect the Way Teachers Teach?	37
3.C	Do You Agree or Disagree That the Process Used by the Kentucky Board of Education To Set Improvement Goals Was Appropriate?.....	41
3.D	Can Your School Reach Proficiency by 2014?.....	42

Program Review and Investigations

3.E	Do You Agree or Disagree That the Consequences to Schools That Fail To Improve Are Appropriate?	43
4.A	Writing Portfolio Audit Process.....	65
5.A	Do You Think Your School/District Can Have a Highly Qualified Teacher in Every Classroom by 2005?	86

Executive Summary

At its July 2002 meeting, the Program Review and Investigations Committee directed staff to review aspects of the Commonwealth Accountability and Testing System (CATS), including writing portfolios. This report presents the results of that review.

Major Conclusions

Self-reported dropout data pose potential problems. Given that dropout information is accepted from school districts by the Kentucky Department of Education without sufficient validation, the accuracy of school, district, and statewide dropout rates is unknown. Other states that have audited self-reported dropout rates have found that rates are typically underreported.

Program Review staff estimated that the CATS assessment cost school districts approximately \$10.6 million for the 2001-2002 school year—about \$16 per enrolled pupil, or \$22 per student tested that year. In fiscal year 2003, the CATS assessment was estimated to cost the state approximately \$10.3 million, again about \$16 per pupil, or \$22 per student tested.

The correlation between individual students' CATS and ACT scores is relatively high, although statewide CATS scores have been increasing and statewide ACT scores have not increased in most recent years.

Just over 50 percent of superintendents and principals surveyed by Program Review staff responded that the quality of education under CATS is better than under the previous system. About a third of teachers said that education is better under CATS; 39 percent said it is about the same as before. Over two-thirds of principals and superintendents responded that CATS affects teaching and learning positively and that they were satisfied with the Core Content for Assessment. Teachers were less supportive than principals and superintendents on the effect of CATS on teaching, learning, and core content, but their most common responses to these questions were still positive.

The evidence from responses to Program Review staff's survey of educators suggests that repetition of questions may affect teaching for some teachers. There is little evidence from the survey, however, to indicate that teachers are "teaching the test" if that is defined as teachers incorporating unreleased past questions into their lessons.

More than one-half of teachers disagreed that the process used to set school improvement goals by the Kentucky Board of Education was appropriate, including 19 percent of respondents who strongly disagreed. Based on their follow-up answers, many teachers appeared to be commenting on the appropriateness of the goals, not just the process used to create them. Many teachers responded that improvement goals are unrealistic because

the system does not measure individual student's progress, but compares different classes. Principals and superintendents responded more positively; one-half of principals and more than 60 percent of superintendents agreed that the process to set school improvement goals was appropriate.

At least 75 percent each of teachers, principals, and superintendents responded to the survey that the accountability weights for reading, math, science, and social studies are about right. Around one-half of teachers responded that the weight for the writing portfolio was too high, with around 30 percent of principals and superintendents indicating the same.

Kentucky educators expressed concern on the survey about the ability to close the disability and income achievement gaps.

Kentucky Department of Education personnel train regional writing consultants, who train writing cluster leaders, who train school-level portfolio scorers. Approximately 90 percent of portfolio scorers responded to the survey that the training provided by the cluster leader was helpful. Teachers reported receiving an average of 4.5 hours of portfolio training annually.

Almost one-half of teachers responded to the survey that some teachers have biases that affect scoring of portfolios.

When asked if their school had enough teachers scoring portfolios, 77 percent of teachers and 89 percent of principals indicated that their school had enough teachers scoring writing portfolios.

The Kentucky Department of Education utilizes a formal audit procedure to increase the probability that portfolios are scored accurately. The portfolio audit is a review of local scoring accuracy and is used to monitor statewide scoring patterns and adjust scores for schools that score portfolios inaccurately. Agreement rates between original scores assigned to writing portfolios and the scores assigned during audits of the writing portfolios do not appear to have increased consistently over recent years.

During the writing portfolio audit, original scores assigned at the school level are revised downward more frequently than upward, indicating that original writing portfolio scores may often be inflated. There are no consequences to writing portfolio scorers or schools for inaccurately scoring portfolios.

Recommendations

- 2.1** KDE should perform specific procedures to verify school districts' reported dropout statistics. As part of the attendance audit, KDE should review the schools' documentation that students coded as transfers are enrolled in other schools. Dropout statistics should be corrected to reflect any inaccuracies found in the audit. In accordance with KRS 158.6458(7), KDE shall report to the Interim Joint Committee on Education on the results of the dropout audit.
- 2.2** In addition to adjusting schools' dropout statistics if errors are found, KDE should consider sanctioning schools that underreport dropout statistics by lowering their scores on the accountability index by an additional amount or by making them ineligible for rewards that year.
- 2.3** KDE should implement a uniform student information system at the state level to track students who transfer, drop out, are expelled, or graduate; and to track the number of students who are excluded from the CATS test.
- 2.4** KDE should create a CATS testing expenditure category in the MUNIS system (the financial software package used for Kentucky school districts) and encourage school districts to utilize this category for all CATS administrative expenses. Any problems in implementing this change should be reported to the Education Committee and the Education Assessment and Accountability Review Subcommittee.
- 4.1** KDE should work with schools and districts to reduce, as much as possible, the practice of teachers scoring their own students' portfolios. KDE should survey teachers to determine how their portfolio scoring training can be improved. KDE should regularly replace benchmark portfolios with new samples. KDE should encourage schools to provide teachers with more opportunities to practice scoring writing portfolios.
- 4.2** KDE should consider implementing a system to track the performance of portfolio scorers.
- 4.3** KDE should consider establishing consequences for schools that have low portfolio audit agreement rates, such as having audit agreement rates reflected in the school's accountability index. KDE should consider reauditing schools that had a high number of scoring inaccuracies the prior year to ensure that scoring accuracy has improved. KDE should also consider increasing the number of schools randomly selected for audits so that the risk of facing consequences would encourage schools to score more carefully.

Chapter 1

An Overview of the Commonwealth Accountability Testing System and This Study

Introduction

In 1998, the General Assembly mandated major changes to Kentucky's education assessment and accountability system.

In 1998, the General Assembly mandated major changes to Kentucky's education assessment and accountability system. The new Commonwealth Accountability Testing System (CATS) included implementing a new test, revising school goals, and revising the consequences to schools based on their performance. Schools are measured and held accountable for how well they are progressing toward meeting the goals included in the state assessment system.

The foundation of the Commonwealth Accountability Testing System (CATS) is the core content, the subject matter identified as essential for all Kentucky students to learn.

The foundation of CATS is the core content, which consists of the subject matter identified by Kentucky educators as essential for all students in the state to learn. The areas of core content are reading, math, science, social studies, arts and humanities, and practical living and vocational skills. These subjects are taught in all public elementary and secondary schools, and teachers are responsible for covering all the core content. The CATS test, administered in the spring of each school year, includes measures of how well students have learned the core content.

Tests should be valid and reliable. Valid means the test measures what it is supposed to measure. For example, the CATS core content test in math is valid if scores indicate how well students have mastered the appropriate math core content. A test is reliable if its results are consistent. Each student answers relatively few questions for the CATS test in each subject area, but because there are multiple forms of the test in each school, the test can be valid at the school level but not at the student level.

Kentucky schools are held accountable for students' test scores in the academic content areas as well as for certain nonacademic factors, which vary by type of school. Elementary, middle, and high schools are accountable for attendance and retention rates. Middle and high schools are accountable for dropout rates. High

schools are also accountable for students' successful transition to adult life.

Under CATS, schools receive sanctions or rewards based on students' performance on the CATS test and nonacademic indicators such as dropout rates.

Under CATS, each school is eligible to receive sanctions or rewards based on its score on its accountability index, which is comprised of students' performance on the CATS test and nonacademic indicators such as dropout rates. If a school meets or exceeds its biennial goal under CATS, it is eligible for monetary rewards. In 2002, the Kentucky Department of Education (KDE) paid out approximately \$22.4 million in school rewards. If a school does not meet its biennial goal under CATS, sanctions could include an audit by the Kentucky Department of Education or dismissal of personnel.

According to national rankings, Kentucky's accountability and assessment system is among the most comprehensive in the country.

Kentucky's education accountability and assessment system is one of the most comprehensive in the United States. *Quality Counts 2003*, the latest of the yearly reports on the states compiled by *Education Week*, ranks Kentucky third on standards and accountability, with a grade of 94 (out of 100). Only Florida and New York received higher grades for the quality of the standards, assessment, and accountability components of their testing systems.

Each state is reorganizing its assessment and accountability system to comply with the federal No Child Left Behind Act.

Each state is reorganizing its assessment and accountability system to comply with the federal No Child Left Behind Act (NCLB), enacted in 2002. Under NCLB, states must define a proficiency level that all students must reach by the 2013-2014 school year. States must expand their testing systems to assess students annually in reading and math in grades 3 to 8 and once in grades 10 to 12. Science must be tested once in elementary, middle, and high school. Other provisions require that states analyze and report test results in new ways, provide supplemental services to students who attend low-performing schools, and help teachers achieve "highly qualified" status.

Description of This Study

How This Study Was Conducted

The Program Review and Investigations Committee voted on July 11, 2002, to have staff study aspects of CATS, including writing portfolios. In conducting the study, Program Review staff reviewed and analyzed data and documentation the Kentucky Department of Education uses to measure educational progress under CATS.

Interviews were conducted with staff of the department, the Kentucky Education Professional Standards Board, the Office of Education Accountability, and the Interim Joint Committee on Education. Staff attended a regional training session for scorers of writing portfolios. Staff also attended two meetings of the National Technical Panel on Assessment and Accountability. Staff surveyed teachers, principals, and superintendents on issues related to CATS assessment and accountability.

Organization of the Report

This report is organized as follows:

The remainder of Chapter 1 summarizes major conclusions from the report, briefly describes the recent history of Kentucky's education system, and describes the CATS test.

Chapter 2 describes Kentucky's accountability system, including the process used to set school improvement goals and the rewards and sanctions applicable to schools based on their performance. Estimates of local and state costs for conducting the CATS assessment are provided.

Chapter 3 summarizes the results of survey questions that called for teachers, principals, and superintendents to evaluate aspects of CATS. This chapter also includes a comparison of CATS and ACT scores.

Chapter 4 describes and evaluates the writing portfolio and on-demand writing components of CATS. Details are provided about the scoring and auditing of portfolios, and the quality of training provided by KDE.

Chapter 5 describes the No Child Left Behind Act, issues regarding Kentucky's compliance with it, and how other states are faring with it.

The Kentucky Department of Education's written response to this report is included as Appendix G.

Major Conclusions

The study's major conclusions are as follows:

The Kentucky Department of Education accepts dropout information from school districts without sufficient validation, so the accuracy of school, district, and statewide dropout rates is unknown.

Program Review staff estimated that for the 2001-2002 school year the CATS assessment cost school districts approximately \$10.6 million—\$16 per enrolled student, or \$22 per student tested that year. In fiscal year 2003, the CATS assessment was estimated to cost the state approximately \$10.3 million, also about \$16 per enrolled student, or \$22 per student tested.

The correlation between individual students' CATS and ACT scores is relatively high, although statewide CATS scores have been increasing while statewide ACT scores have not increased in most recent years.

More than 50 percent of principals and superintendents responded to Program Review staff's survey that the quality of education is better under CATS compared to the previous system. About a third of teachers said that education is better under CATS.

There is little evidence from the survey to indicate that teachers are "teaching the test" if that is defined as teachers incorporating unreleased past questions into their lessons.

More than half the teachers disagreed that the process used to set school improvement goals by the Kentucky Board of Education was appropriate. Many teachers responded that improvement goals are unrealistic because the system does not measure individual students' progress, but compares different classes.

At least 75 percent each of teachers, principals, and superintendents responded to the survey that the accountability weights for reading, math, science, and social studies are about right. Around one-half of teachers responded that the weight for the writing portfolio was too high, with around 30 percent of principals and superintendents indicating the same.

Almost one-half of teachers responded to the survey that some teachers have biases that affect scoring of writing portfolios.

Agreement rates between original scores assigned to writing portfolios and the scores assigned during audits of the writing portfolios do not appear to have increased consistently over recent years. During the writing portfolio audit, original scores assigned at the school level are revised downward more frequently than upward, indicating that original writing portfolio scores often may be inflated. There are no consequences to writing portfolio scorers or schools for inaccurately scoring portfolios.

The Kentucky Education Reform Act

In 1989, the Kentucky Supreme Court ruled that the state's school system was unconstitutional and called for a system based on efficiency and defined by adequacy and equity.

In 1989, the Kentucky Supreme Court ruled that the state's school system was unconstitutional. The court's decision called for a system based on efficiency and defined by adequacy and equity. In response, the 1990 General Assembly enacted House Bill 940, known as the Kentucky Education Reform Act (KERA). KERA, deemed one of the most comprehensive educational reform acts in the United States, called for systematic changes in assessment, curriculum, governance, and finance. Through KERA, the General Assembly mandated the creation and implementation of a statewide performance-based student assessment program and school accountability system. As the timeline in Table 1.1 illustrates, however, policies had already been established in Kentucky that addressed issues of student assessment and school accountability.

Goals of the Kentucky Education Reform Act

The Kentucky Education Reform Act established school goals, student learning goals, and academic expectations.

Following KERA, the General Assembly and the Kentucky Board of Education established the standards Kentucky's public schools should meet through the creation of goals for schools, student learning goals, and academic expectations. KRS 158.6451 states that Kentucky's public elementary and secondary schools shall:

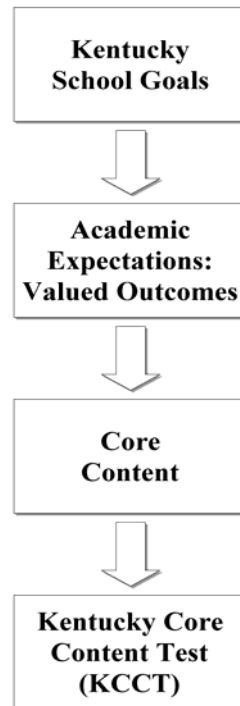
- Expect a high level of achievement of all students;
- Develop students' abilities in six cognitive areas;
- Increase school attendance levels;
- Reduce dropout and retention rates;
- Reduce physical and mental health barriers to learning; and
- Increase the proportion of students who make a successful transition to work, postsecondary education, and the military.

Table 1.1
Timeline of Events Prior to KERA

1978. The General Assembly enacts the Education Improvement Act, which mandates statewide annual achievement testing in grades 3, 5, 7, and 10 and requires the purchase or development of appropriate tests (codified as KRS 158.650 – KRS 158.730). 1979. Statewide achievement testing begins with a commercially available norm-referenced test: Comprehensive Test of Basic Skills, Third Edition (CTBS-3). 1984. The General Assembly amends KRS 158.690 to require districts to publish annual performance reports in newspapers. 1985. The Council for Better Education sues state officials, alleging Kentucky’s educational system is unconstitutional. The Kentucky Department of Education establishes more than 600 essential skills for students. 1985 to 1988. CTBS-3 is discontinued and is replaced by the customized Kentucky Essential Skills Test. 1988. The General Assembly amends KRS 158.700 to require the purchase of a standardized, nationally normed test to administer to students, removing the discretion to develop customized tests. 1989. Governor Wallace Wilkinson creates the Council on School Performance Standards to study educational goals and learning assessment. The council presents its report and recommendations three months after the supreme court decision. 1989 to 1990. The Kentucky Essential Skills Test is replaced by a commercially available, norm-referenced test: Comprehensive Test of Basic Skills, Fourth Edition (CTBS-4).

The relationship between the different goals, expectations, and testing of students is illustrated in Figure 1.A. Based on the school goals, in 1994 the Kentucky Board of Education approved 57 valued outcomes known as the academic expectations—measurable values to gauge how well schools are meeting their goals. Based on the academic expectations, the core content was defined to give educators information on how to cover the material students are expected to learn to meet the academic expectations. Finally, the Kentucky Core Content Test (KCCT), which makes up almost all the CATS test, measures how well students have learned the core content.

Figure 1.A
How School Goals Relate to
the Kentucky Core Content Test



For example, one goal calls for schools to develop students' abilities to apply core concepts and principles from mathematics. Academic expectation 2.8 calls for students to understand various mathematical procedures and use them appropriately and accurately. As covered in the core content, students should be able to use standard units to measure volume of rectangular prisms, liquid capacity, money, time, and temperature. This material is then assessed on a core content test. The example below from the 1999 Grade-5 KCCT tests this particular core content.

Anton has saved 37 nickels. What is the total value of the nickels he saved?

A. \$1.35
B. \$1.55
C. \$1.85
D. \$10.35

The Commonwealth Accountability Testing System

As the learning goals and academic expectations were being created, the Kentucky Instructional Results System (KIRIS) was established as the state assessment system. KIRIS was developed

in 1992 to measure progress toward the goals established under KERA. Criticism of KIRIS became widespread. In 1994, a panel of measurement specialists was appointed by the Office of Educational Accountability and the General Assembly to investigate the technical quality of KIRIS. According to the panel's report, the test frameworks did not communicate clearly what students were expected to know, the test scores reported for schools were not adequately reliable for accountability purposes, the student performance standards lacked standardization, and writing portfolio scores were not reliable (Hambleton et al., 1995). The Task Force on Public Education, established in 1996 to investigate the effectiveness of KIRIS, recommended changes in Kentucky's system of assessment and accountability.

In 1998, HB 53 amended KRS 158.6453, creating the Commonwealth Accountability Testing System.

In 1998, HB 53 amended KRS 158.6453, replacing KIRIS with the Commonwealth Accountability Testing System. The legislation mandated that CATS include customized or commercially available norm-referenced tests that measure the core content for assessment and that also provide valid and reliable results for individual students. The legislation also required that teachers play a significant role in the design of the new assessments. The legal basis for the components of CATS is reviewed in Appendix A, which summarizes relevant Kentucky statutes and regulations.

CATS was designed to address the perceived limitations of KIRIS in the following ways:

- Scores are to be used for school accountability;
- Teachers are to be extensively involved in designing and scoring the test;
- Results of a nationally norm-referenced test should be used to compare Kentucky schools to those in other states;
- The new format should reduce testing times for schools and students; and
- Test results are to be reported to schools and districts in a more timely fashion.

The CATS Test

Under CATS, different tests and types of test items are utilized to measure student performance in several subject areas.

CATS uses different tests and types of test items to measure student performance in several subject areas. CATS uses both norm-referenced and criterion-referenced tests. The norm-referenced test is used to compare Kentucky's students to those in other states. The criterion-referenced test is used to measure how well students have learned the core content. Multiple-choice and open-response questions and portfolios are used to assess students'

knowledge and abilities. Appendix B contains samples of questions by type, grade level, and subject.

Norm-referenced and Criterion-referenced Tests

Under CATS, there are two different tests administered to students: the Kentucky Core Content Test and the Comprehensive Test of Basic Skills, Fifth Edition (CTBS-5). The KCCT is the criterion-referenced test administered to students in grades 4, 5, 7, 8, 10, 11, and 12. The KCCT tests students in reading, math, science, social studies, arts and humanities, practical living and vocational studies, and writing. The CTBS-5 is the nationally norm-referenced test that assesses students in reading, math, and language arts in grades three, six, and nine.

A norm-referenced test is designed to assess achievement differences between and among students. A criterion-referenced test is designed to evaluate what test takers can do and what they know.

A norm-referenced test such as the CTBS-5 is designed to assess achievement differences between students. The content of a norm-referenced test is selected based on the desire to rank students from high to low achievers. The main reason for using the CTBS-5 norm-referenced test is to measure how Kentucky's students are progressing compared to a group established as the norm in 1996. Comparisons of percentile rankings over time are also possible with other states that are using the same form of the test. The validity of a norm-referenced test depends on whether the content tested matches the knowledge and skills expected of students in a particular school system. Each school's students' performance on the CTBS-5 counts for 5 percent of the index through which the school is held accountable.

Based on their performance on core content tests, Kentucky students are classified as novice, apprentice, proficient, or distinguished.

A criterion-referenced test is designed to evaluate what students can do and what they know based on a standard of performance. It does not, however, evaluate referenced test how students compare to others. A criterion-referenced test assesses how well students are doing relative to a predetermined performance level on a specified set of educational goals or outcomes. The purpose of the Kentucky Core Content Test is to measure how well students have mastered the core content. Based on their performances on the tests that make up the KCCT, students are divided into four categories: novice, apprentice, proficient, and distinguished. The performance of students on the core content tests makes up the majority of a school's accountability index.

It is easier to ensure the match to expected skills with a criterion-referenced test because the curriculum in the schools should match the content on the test. In Kentucky, the KCCT measures students'

mastery of the core content. If the content of the test matches the content that is considered important to learn, the criterion-referenced test provides more information than the norm-referenced test about how much of the valued content has been learned through instruction.

For example, on a criterion-referenced math test, 11th-grade proficiency criteria might be to solve algebra, geometry, or trigonometry problems. The test will pose such problems and award a score of “11th-grade proficiency” only if the test taker succeeds on these particular problems. The results of a norm-referenced math test might yield a percentile rank score of 60, which communicates that the student performed as well or better than 60 percent of the students in the norm group.

Subjects Tested Under CATS

Under CATS, students are tested in the subject areas of reading, writing, mathematics, science, social studies, arts and humanities, and practical living and vocational studies.

Under CATS, for accountability purposes students are tested in the subject areas of reading, writing, mathematics, science, social studies, arts and humanities, and practical living and vocational studies. Table 1.2 illustrates the grades and subject areas in which students are tested under CATS. Schools are held directly accountable for students’ scores in all the subject areas that are tested.

Table 1.2
Subjects Tested Under CATS

Subject:	Grade:	3	4	5	6	7	8	9	10	11	12
Reading			•			•			•		
Math				•			•			•	
Science			•			•				•	
Social Studies				•			•			•	
Arts and Humanities				•			•			•	
Practical Living and Vocational Studies				•			•		•		
On-demand Writing			•			•					•
Writing Portfolio			•			•					•
CTBS-5 Norm-referenced Test*		•			•			•			

*CTBS-5 tests students in reading, math, and language arts.

Source: Kentucky Department of Education.

Types of Test Items

Kentucky uses multiple-choice and open-response questions and writing portfolios.

Kentucky uses multiple-choice, open-response, and portfolio-item questions in accordance with KRS 158.6453. Each item type seeks to identify a student's level of thinking and knowledge.

Multiple-choice questions ask for the best answer from a list of alternatives.

Multiple Choice. Multiple-choice tests include questions in which two or more answers are offered as alternative responses for each item. Multiple-choice questions ask for the best answer to a question or an incomplete statement.

Multiple-choice questions cover a broad range of content and can be scored objectively and quickly, but provide no measure of students' writing abilities. Below is an example of a multiple-choice question from the Grade 4, Spring 1999 CATS test.

Soil is made up MOSTLY of
A. helium.
B. weathered rock.
C. water.
D. air.

Open-response questions require written answers by students.

Open Response. An open-response question requires students to answer questions in writing. Responses can be brief or extensive. Open-response questions test for recall and the ability to apply knowledge of a subject to questions about the subject.

Open-response questions differ in how much freedom is permitted to the student in making a response. Restricted responses require brief and precise answers to specific questions. Extended responses reflect more comprehensive questions that allow greater freedom in structuring a response. Compared to other states, Kentucky utilizes many open-response questions on the KCCT. Each of the six test forms in reading, math, science, and social studies contains six open-response questions. Each of the 12 test forms in arts and humanities and practical living and vocational studies contains two open-response questions.

Advantages to using open-response questions are that students are less likely to guess and questions are easier to construct. Such questions allow students to organize knowledge, express opinions, and show originality. Because of the time necessary to answer them, open-response questions may limit the amount of material tested. Compared to multiple-choice questions, they are also more

difficult and time consuming to score. Below is an example of an open-response question from the Grade 7, Spring 1999 CATS test.

11. Plants and animals rely on one another for the production of oxygen and carbon dioxide.
- a. Describe this relationship.
 - b. Use a diagram or flow chart to illustrate your description of this relationship.

A portfolio is a collection of student work that is usually drawn from students' classroom assignments.

Portfolios. A portfolio is a collection of student work that is usually drawn from students' classroom assignments. Kentucky and Vermont are the only two states that use portfolios in their statewide testing systems.

Portfolios are used to assess students' writing, documentation, critical thinking, and problem solving skills. Portfolios can allow students to present the totality of their learning and allow students to become more active participants in the evaluation process. For example, in Kentucky students work closely with their teachers in selecting and revising writing pieces for their portfolios. The disadvantages of using portfolios include the limit on the amount of material tested, potentially unreliable scoring, and the time necessary to score them.

Chapter 2

Kentucky's Education Accountability System

Each public school in Kentucky has its own growth chart, detailing the scores it must attain to meet its two-year goals and the long-term goal of proficiency by 2014.

Kentucky's long-term accountability model was implemented in 2002. The Kentucky Board of Education established the accountability system to compare individual school progress in two-year cycles using academic and nonacademic indicators. Each public school in Kentucky has its own growth chart, detailing the scores it must attain to meet its two-year goals and the long-term goal of proficiency by 2014. Proficiency means that a school scores at least 100 on a 140-point scale, meaning that almost all students must score at least proficient.

Schools' Two-year Goals

Calculation of Goals

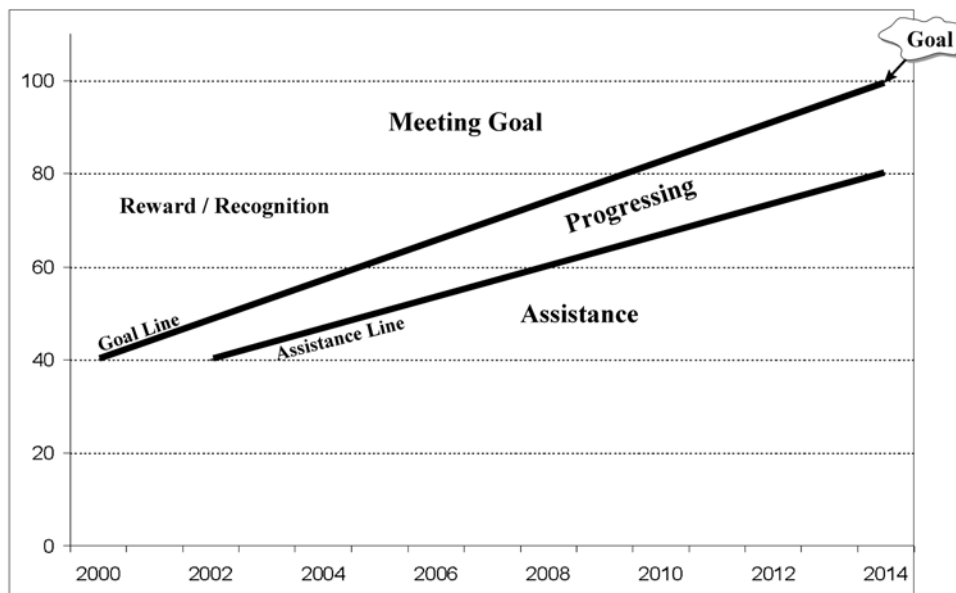
Each school's growth chart is formulated as if the school would reach the long-term goal of 100—defined as proficiency on a 140-point scale—in seven equal steps, with each step taking two school years. Reaching or surpassing the goal of 100 earlier would be better, obviously. A school with a baseline near 100 in 2000 has smaller biennial steps. Schools that began with baselines further below 100 must attain greater biennial improvements. Figure 2.A is an example of what a school's growth chart might look like.

The Kentucky Board of Education is responsible for the details of Kentucky's accountability system, including the process for calculating biennial goals for schools.

The Kentucky Board of Education is responsible for the details of Kentucky's accountability system, including the process for calculating biennial goals for schools. The biennial school goals were calculated as follows:

1. Each school's 1999 and 2000 test results were averaged to determine its starting point or baseline.
2. The baseline was subtracted from 100 to determine how much the school needed to improve to reach the long-term goal of 100 in 2014.

Figure 2.A
A School's Long-term Growth Chart



Source: Kentucky Department of Education.

3. There are seven two-year periods between 2014 and 2000, so the amount the school must improve was divided by seven to determine how much improvement is needed every two years.
4. Each of the resulting two-year goals was lowered to reflect the test's margin of error. The margin of error is represented by the width of the goal line and the assistance line as seen in Figure 2.A.¹

As shown in the example below, if a school's baseline index score is 51, then 51 is subtracted from 100 to determine how much improvement is needed to meet the goal of 100 in 2014. The difference between 100 and 51 (49) is divided by 7 to yield 7, the required improvement for each two-year period until 2014. The margin of error, calculated by the Kentucky Department of Education (KDE) based on school size and number of students tested, is subtracted from each two-year goal. For this example, the margin of error is 1.

¹ The margin of error accounts for fluctuations in scores due to chance occurrences that would affect scores, such as students' health at the time of assessment. The fewer the students in a school, the larger its margin of error. KDE, using 1999 and 2000 test data, calculated each school's margin of error.

2000: Baseline = 51	Goal minus margin of error
2002: Goal = $51 + 7 = 58$	$58 - 1 = 57$
2004: Goal = $58 + 7 = 65$	$65 - 1 = 64$
2006: Goal = $65 + 7 = 72$	$72 - 1 = 71$
2008: Goal = $72 + 7 = 79$	$79 - 1 = 78$
2010: Goal = $79 + 7 = 86$	$86 - 1 = 85$
2012: Goal = $86 + 7 = 93$	$93 - 1 = 92$
2014: Goal = $93 + 7 = 100$	$100 - 1 = 99$

Rewards and Assistance

Based on its index score, a school is classified as meeting its goal, progressing, or needing assistance.

The goal line is a straight line that begins at a school's baseline and ends in 2014 with a score of 100 on the accountability index (see Figure 2.A). A school that is meeting its goal may qualify for financial rewards and recognition. A school that falls below the goal line, but above the assistance line is "progressing" and may also qualify for rewards. The assistance line is a straight line that begins in 2002 at the baseline and ends in 2014 with a score of 80. "Needs assistance" schools are those that fall below the assistance line. The local school district and KDE are to assist these schools in improving.

For each two-year accountability cycle, each school must progressively reduce the percentage of students who score in the novice level so that no more than 5 percent of its students score novice by 2014. Schools also must have a two-year dropout rate below 5.3 percent, or a two-year rate that is 6 percent or lower and has declined at least one-half of 1 percent since the previous biennium. The novice requirement is applicable to all schools; the dropout requirement is applicable to middle and high schools.

In accordance with KRS 158.6455 and 703 KAR 5:020, schools qualify for financial rewards in an accountability cycle if they meet the dropout and novice reduction targets and 1) meet or exceed their goals; or 2) score in the progressing area with an index greater than they had in the last biennium; or 3) pass through any of the recognition points 55, 66, 77, 88, or 100 for the first time if the recognition point was above the school's baseline. Schools that score in the top 5 percent of all schools, meet dropout and novice reduction targets, meet or exceed the fourth recognition point, and did not decline in the previous two biennia are designated Commonwealth pace-setter schools and also receive rewards if not already receiving rewards under another provision.

It is not uncommon for a school to meet its accountability index goal, but not receive rewards because it failed to reduce its percentage of novices or dropouts.

It is not uncommon for a school to meet its accountability index goal, but not receive rewards because it failed to reduce its percentage of novices or dropouts. According to the CATS scores for the 2000-02 biennium, all but 90 of 1,182 Kentucky schools bettered their scores. A third of the improving schools did not receive monetary rewards because either not enough of their students had progressed from the novice category or the school did not meet the dropout requirement.

Kentucky's School Accountability Index

A school's progress is measured by its accountability index, which includes academic and nonacademic components. The components of the nonacademic index vary by whether the school is an elementary, middle, or high school.

A school's progress is based on its accountability index score, which is calculated based on the results of tests given in the preceding spring, plus certain nonacademic components. Schools receive their annual accountability scores in October. The index consists of the following components (each component's weight in the index is shown in Table 2.1 on the next page):

- The CTBS-5 norm-referenced test, a multiple-choice test that allows comparisons nationally between Kentucky students and their peers in language arts, reading, and math;
- The Kentucky Core Content Test, a mixture of multiple-choice and open-response questions in reading, science, math, social studies, arts and humanities, and practical living and vocational studies;
- The writing portfolio, a collection of a student's best writing over time;
- On-demand writing, writing tests that measure skills developed from writing instruction; and
- The nonacademic index. The nonacademic index consists of:
 - Attendance rate—the average percentage of students who attend school daily;
 - Retention rate—the percentage who are retained or held back in a grade;
 - Dropout rate—the percentage who drop out of school; and
 - Transition to adult life—the percentage of high school students who make a successful transition to adult life by getting a job, entering the military, or continuing their education.

The dropout rate is only applicable in the middle and high school levels and is given more weight at the high school level. Transition to adult life is only applicable at the high school level. Attendance and retention rates are applicable at all levels. School districts report values for the nonacademic index to KDE.

Table 2.1 illustrates the weights of the different academic and nonacademic components in a school's accountability index by grade level. The weights of the academic and nonacademic components in the accountability index were assigned by the School Curriculum, Assessment and Accountability Council and written into regulation in August 1999.

Table 2.1
Percentage Weights of Components in the
School Accountability Index by Grade Level

Component	Grade Level		
	Elementary	Middle	High
Reading	19.00%	14.25%	14.25%
Mathematics	19.00	14.25	14.25
Science	14.25	14.25	14.25
Social Studies	14.25	14.25	14.25
On-demand Writing	2.85	2.85	2.85
Writing Portfolio	11.40	11.40	11.40
Arts and Humanities	4.75	7.125	7.125
Practical Living and Vocational Studies	4.75	7.125	7.125
Attendance Rate	3.80	3.80	1.90
Retention Rate	0.95	3.80	0.48
Dropout Rate	—	1.90	3.56
Transition to Adult Life	—	—	3.56
National Norm-referenced Test (CTBS-5)	5.00	5.00	5.00
Total	100%	100%	100%

Source: 703 KAR 5:020.

CATS Accountability and School Rewards

Based on statute and regulation, money is paid from the Kentucky Successful Schools Trust Fund to schools qualifying for rewards.

As required by 703 KAR 5:020, 703 KAR 5:040, and KRS 158.6455, a sum of money is paid from the Kentucky Successful Schools Trust Fund to schools qualifying for rewards. The Kentucky Successful Schools Trust Fund consists of money the General Assembly sets aside to reward schools.

The Kentucky Department of Education determines a reward amount consistent with regulatory requirements to be distributed to every school qualifying for rewards based on academic and nonacademic components. The Kentucky Board of Education then must authorize the distribution of the reward money.

Rewards are distributed to schools based on the number of full-time certified staff employed in the school on the last working day of the year of the reward. The total amount of rewards distributed to schools and school districts cannot exceed 1.75 percent of the amount of funds paid to certified personnel (teachers and administrators) during the last year of the accountability cycle. A reward share cannot exceed \$2,000.²

Under Section 9 of 703 KAR 5:020, a school classified as meeting goal earns three shares of rewards, not exceeding \$6,000. A school classified as progressing receives a one-half share of rewards, not exceeding \$1,000. For example, if a reward share equals \$500 and a school has 50 full-time certified staff and is eligible for three shares of rewards (50 x 3 x \$500), the school's total reward is \$75,000.

In 2002, the Kentucky Department of Education paid out more than \$22 million in school rewards. School councils determine the use of reward money earned by a school.

Once the amounts of rewards are calculated, the Kentucky Department of Education distributes the funds to local school districts, which disburse the funds to the reward school(s). All reward funds are accounted for at the local school district level. In 2002, the Kentucky Department of Education paid out more than \$22 million in school rewards. A total of 736 schools across the state received reward money. School councils determine the use of reward money earned by a school. Schools often use reward money to purchase materials, to provide training, or to provide bonuses to teachers.

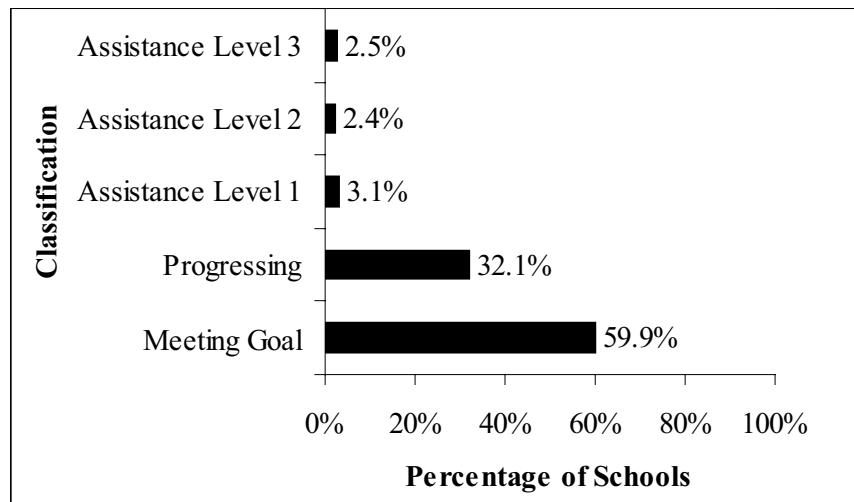
“Needs Assistance” Classification System

In 2002, more than 80 percent of Kentucky schools were classified as meeting goal or progressing.

As shown in Figure 2.B, in 2002 more than 80 percent of Kentucky schools were classified as meeting goal or progressing. A school that falls below its assistance line is classified as needs assistance. All schools falling into the assistance classification are ranked from highest to lowest according to the school's combined biennial accountability index. The group of schools is then divided into thirds. The top third are designated Level 1 schools, the middle third Level 2, and the bottom third Level 3, with consequences and sanctions varying by level. Details are provided in Appendix C.

² A reward share is the unit of money to be distributed to schools, and is determined by the total amount of money available for rewards in a biennium and the total number of shares to be awarded.

Figure 2.B
2002 Classification of Schools



Source: Compiled by Program Review staff from KDE data.

Dropout Rates and CATS

Schools that meet their accountability index but do not have a dropout rate lower than 5.3 percent will not receive reward money.

Schools that meet their accountability index but do not have a dropout rate lower than 5.3 percent will not receive reward money. In 1999, the School Curriculum, Assessment and Accountability Council (SCAAC) recommended 5.3 percent as the threshold because it was the statewide average high school dropout rate at that time. SCAAC, consisting of 17 voting members appointed by the governor, was created by HB 53 to study, review, and make recommendations concerning Kentucky's system of setting academic standards, assessing learning, holding schools accountable for learning, and assisting schools to improve their performance.³ By July 2006, no school is to have an annual dropout rate exceeding 5 percent (KRS 158.145).

HB 53 changed the formula used to calculate dropout rates to align it with the formula used by the National Center for Education Statistics.

The Kentucky Board of Education changed the formula used to calculate dropout rates to align it with the formula used by the National Center for Education Statistics. The primary change was the year and grade to which dropouts would be attributed. Prior to the 1998-1999 school year, a student who failed to return to school by October 1 was counted as a dropout for the previous school year and the grade just completed. Under the current formula, a student who fails to return to school by October 1 is considered a dropout

³ The members of SCAAC are two parents, two teachers, two principals, two superintendents, two local school board members, two district assessment coordinators, two employers representing private business, two university professors, and one at-large member.

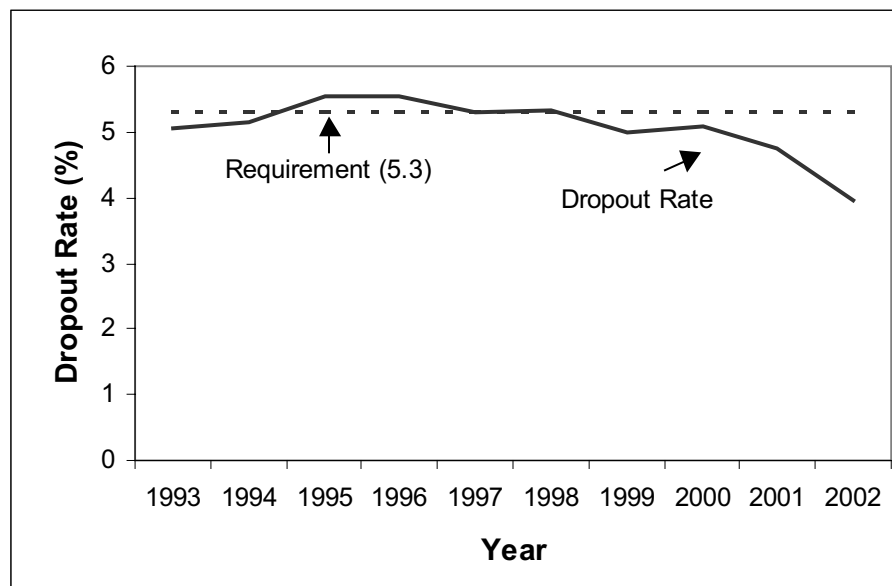
for the grade and year in which he or she did not return. Program Review staff calculated the statewide high school dropout rates for the school years 1998-1999 to 2001-2002 using the old and new formulas. Differences were insignificant.

KDE Reports That the Dropout Rate Has Declined

The state high school dropout average has fluctuated over the past decade.

The state high school dropout average has fluctuated over the past decade, as seen in Figure 2.C. The dropout rate peaked in 1995 and 1996, with a statewide average of 5.53 percent. The rate has not declined in every year since, but the general trend is toward lower dropout rates. In 2002, the rate declined to 3.97 percent.

Figure 2.C
State Dropout Requirement and State Dropout Rate, 1993 to 2002



Source: Compiled by Program Review staff from KDE data.

Reporting Procedures for Dropouts

Local school districts report dropout rate data to KDE by November 1 each year.

Local school districts report dropout rate data to KDE electronically by November 1 each year. Some districts choose to collect the dropout data from the schools themselves, others let school personnel handle the compilation of dropout data.

When a student leaves a school, the student must be classified as a dropout, a transfer, another type of withdrawal (such as an

expulsion), or a graduate. Students who drop out are supposed to complete the Student Dropout Questionnaire documenting their reason for leaving school. School guidance counselors complete this questionnaire during a one-hour counseling session with students and parents and the questionnaire is retained in the student's permanent file.

School personnel are responsible for completing worksheets stating the total number of dropouts during the school year. Data also is broken down by race, ethnicity, and gender. These worksheets are sent to the district central office. Once data for all schools in the district is entered, the district assessment coordinator must certify that the data is complete and correct. The data is then sent to KDE via the Internet.

Self-reported Dropout Data Pose Potential Problems

KDE requires schools to use 16 different codes to differentiate between the types of withdrawals. Consider the following examples:

- A public school student who transfers to a nonpublic school in the same district is coded W03.
- A code of W05 is used for a student who has moved out of a public school district and for whom a request for student records has been received from another school district.
- A student who is 16 but not yet 18 years of age and has dropped out is coded W06.
- A code of W11 is used for a student who has been expelled for behavioral reasons and has withdrawn to a regional alternative facility not run by the expelling local school district.

Errors made in coding withdrawals could affect a school's self-reported dropout rate. Given that dropout information is accepted from school districts by KDE without sufficient validation, the accuracy of school, district, and statewide dropout rates is unknown.

Errors among the codes could affect a school's self-reported dropout rate. For example, if a student moves out of a district, the prior school will record the withdrawal as a transfer if another district requests the student's records. However, if the student never actually attends school in the new district, the student should be recorded as a dropout rather than a transfer.

KDE attendance auditors are directed to verify whether proper documentation is available to confirm student withdrawals. However, specific procedures are not directed toward verifying the number of dropouts and other withdrawals to detect coding errors.

National studies and dropout audits performed in other states have found consistently that actual dropout rates are higher than the reported rates.

When applicable, a school must have a dropout rate of less than 5.3 percent to receive reward money. If school districts underreport the number of dropouts to KDE, schools that do not meet the dropout requirement may be receiving school reward money. Given that KDE accepts dropout information from school districts without sufficient validation, the accuracy of school, district, and statewide dropout rates are unknown. National studies and dropout audits performed in other states have found consistently that actual dropout rates are higher than the reported rates.

A statewide audit in Texas, done by the Office of the State Auditor, found that the actual dropout rate in 1994 was twice the figure reported by the Texas Education Agency. The Texas Auditor's Office compared 1994 statewide high school enrollment with the prior year and found that more than 113,000 students enrolled in 1993 were unaccounted for in 1994. These missing students were not reported as enrolled, graduated, or dropped out. The analysis concluded that 42 percent of the missing students should have been reported as 1994 dropouts, raising the statewide dropout rate from 2.6 percent to 5.7 percent.

Between 1999 and 2003, the New York Education Department Office of Audit Services concluded that several school districts in the state submitted inaccurate dropout data. For example, one district reported a dropout rate of 1.2 percent for the school year 1999-2000. After reviewing records for students who had been coded as withdrawn, the auditors determined that the actual dropout rate was 6.7 percent. In 2002, an audit of another district in New York concluded that its dropout rate was actually 11.1 percent, almost three times higher than the 4.2 percent rate calculated by the district.

A study done for the Manhattan Institute (Greene, 2002) found that Washington's 2001 statewide dropout rate was 33 percent, nearly twice as high as the statewide dropout rate of 18 percent reported by the Washington Office of Superintendent of Public Instruction.

On a national level, a recently released report from a study conducted by the Center for Labor Market Studies at Northeastern University concluded that the nation's high school dropout rate may be as high as 30 percent, almost three times higher than the federal government estimates (Sum and Harrington, 2003). According to the study, a more accurate way to calculate high school graduation rates is to compare the annual number of diplomas awarded by public and private high schools to the number of 17- or 18-year-olds in the United States. Using this

method, which has been adopted by the U.S. Department of Education and many researchers, the nation's high school graduation rate was only 70 to 71 percent in recent years. Based on this calculation, the national dropout rate has stayed about the same for the past 20 years.

In sum, there is evidence that true dropout rates may be higher than reported rates. Given that Kentucky schools are supposed to reduce the percentages of students who drop out, it is critical that the reported figures be accurate. There are complementary approaches that could be taken to inspire more confidence in Kentucky's reported dropout rates. In the short term, KDE could make more of an effort to verify schools' dropout data. Schools should also be held more accountable for the accuracy of the dropout information reported to KDE.

Recommendation 2.1

KDE should perform specific procedures to verify school districts' reported dropout statistics. As part of the attendance audit, KDE should review the schools' documentation that students coded as transfers are enrolled in other schools. Dropout statistics should be corrected to reflect any inaccuracies found in the audit. In accordance with KRS 158.6458(7), KDE shall report to the Interim Joint Committee on Education on the results of the dropout audit.

Recommendation 2.2

In addition to adjusting schools' dropout statistics if errors are found, KDE should consider sanctioning schools that underreport dropout statistics by lowering their scores on the accountability index by an additional amount or by making them ineligible for rewards that year.

The Importance of a Statewide Student Information System

A longer-term solution would be KDE's adoption of a statewide student information system. An adequate student-level data system would allow KDE to better track students who transfer, drop out, or graduate and also would aid the department in monitoring the number of students who are excluded from the CATS test.

In recent years, many state education agencies have built the capacity to collect and manage sizable quantities of information regarding schools and districts, and to match individual student records over time and across databases. According to *Quality Counts 2003*, 17 states currently utilize a statewide student information system to collect student-level information on enrollment, transfers, untested students, course-completion, test results, and graduation and dropout data. California, Indiana, Maine, Michigan, and North Carolina have statewide student identification systems currently under development.

A statewide student information system would allow KDE to account for students better.

KDE does not collect individual-level data on transfers, dropouts, or graduates. A statewide student information system would allow KDE to account for students better. For example, if a student is marked as transferring to another high school at the end of a school year, through the use of a student information system, KDE would be able to check if that student showed up at the new school the next year. The data system would allow KDE to hold schools and districts accountable for accurately coding students as transfers, dropouts, or graduates.

The No Child Left Behind Act requires Kentucky to keep track of the number and percentage of students who are excluded from the CATS test each year. Through a student information system, KDE would be able to maintain individual student records and match those records to a separate enrollment database. This would make it possible to account for students from different groups, such as students with disabilities or limited English proficiency, and to check for unusually high rates of test exclusion. Including all students is important under the No Child Left Behind Act, and implementing a statewide student information system would allow KDE to monitor schools and districts to ensure that all students who can be tested are tested.

Recommendation 2.3

KDE should implement a uniform student information system at the state level to track students who transfer, drop out, are expelled, or graduate; and to track the number of students who are excluded from the CATS test.

Graduation Rates

To comply with the No Child Left Behind Act, Kentucky and 19 other states will be using the National Center for Education Statistics' (NCES) completion-rate formula to calculate high school graduation rates.

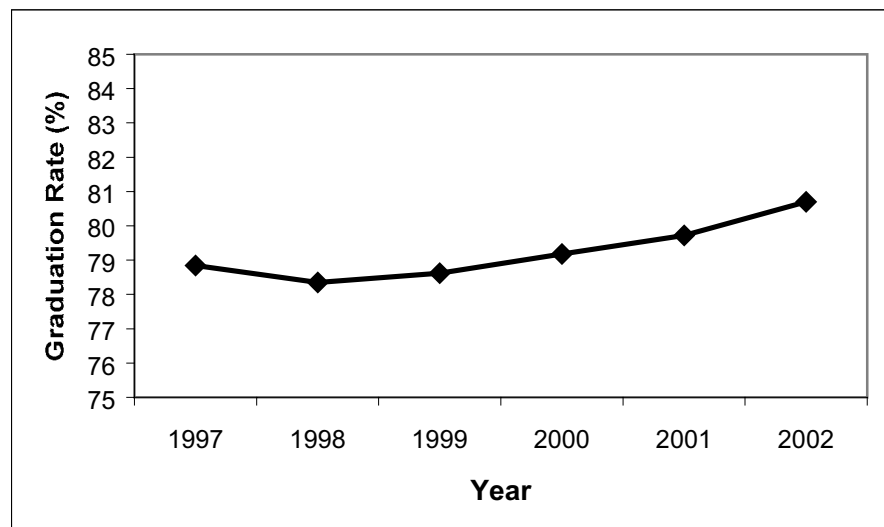
Rather than measuring the number of students who drop out of school, the graduation rate focuses on the rate at which students complete their education through high school. To comply with the No Child Left Behind Act, Kentucky and 19 other states will be using the National Center for Education Statistics' (NCES) completion-rate formula to calculate high school graduation rates. The NCES formula measures the attrition of students over four years from one graduating class by subtracting dropouts from the number of students who complete high school. Using the formula, the 2002 completion rate would be calculated like this:

$$\frac{2002 \text{ Completers}}{(2002 \text{ Completers} + 2002 \text{ Grade-12 Dropouts} + 2001 \text{ Grade-11 Dropouts} + 2000 \text{ Grade-10 Dropouts} + 1999 \text{ Grade-9 Dropouts})}$$

The NCES formula is dependent on dropout rates, so the accuracy of graduation rates will only be as good as the accuracy of dropout reporting. Since a low of 78.4 percent in 1998, the statewide graduation rate has increased each year to reach 80.7 percent in 2002.

The NCES formula is dependent on dropout rates, so the accuracy of graduation rates will only be as good as the accuracy of dropout reporting. Program Review staff utilized KDE data to calculate statewide graduation rates using the NCES formula. Only students who received a regular high school diploma were included in the formula. Figure 2.D displays the statewide graduation rate based on the NCES formula for the past five years. Since a low of 78.4 percent in 1998, the rate has increased each year to reach 80.7 percent in 2002.

Figure 2.D
Statewide High School Graduation Rate Based on the NCES Formula, 1997 to 2002



Source: Compiled by Program Review staff from KDE data.

Graduation Formulas Yield Different Results

A 2003 Urban Institute study indicates that using the NCES graduation formula results in relatively high graduation rates compared to alternative formulas.

A 2003 Urban Institute study (Chaplin) indicates that using the NCES graduation formula results in higher graduation rates than alternative formulas. The report stated that the best evidence available indicates a move away from reliance on graduation rates using dropout data, which may be substantially underreported.

The authors of the Urban Institute report calculated graduation rates using three different formulas: the NCES formula, a formula the authors developed themselves called the Cumulative Promotion Index (CPI), and the Adjusted Completion Ratio (ACR) formula.

The CPI graduation formula estimates the probability that a student entering the ninth grade will complete high school on time with a regular diploma. It does so by multiplying the proportion of 12th-grade students who earn diplomas in a given year with the percent of students in grades 9, 10, and 11 who are promoted to the next grade that same year.

Many states use the ACR graduation formula. This approach compares the number of graduates in a given year with the size of the ninth-grade class four years earlier, adjusting for changes in district enrollment due to transfers.

Kentucky's high school graduation rate for the 2000-2001 school year varied from 64 percent to 80 percent depending on the graduation formula used.

Program Review staff calculated Kentucky's high school graduation rate for the 2000-2001 school year using the same formulas used in the Urban Institute report. The CPI formula produced a statewide graduation rate of 64 percent, the ACR rate produced a graduation rate of 70 percent, and the NCES formula, which Kentucky plans to use, produced a graduation rate of 80 percent.

The Financial Cost of the CATS Assessment

Program Review staff estimated that the local school districts' total costs for CATS assessment were about \$10.6 million for the 2001-2002 school year. State costs for FY 2003 were about \$10.3 million.

The Program Review and Investigations Committee directed that staff estimate the financial cost of the CATS assessment. Staff estimated that the local cost of CATS assessment was about \$10.6 million for the 2001-2002 school year. State costs for fiscal year 2003 were about \$10.3 million. These numbers include only the monetary costs that can be attributed to an assessment task and not to any other activity. For example, there is no cost included for the use of school facilities, such as classrooms, which were not built and are not maintained because of testing requirements.

The cost of the assessment is calculated using information from KDE, the state Management Reporting Database, MUNIS (the financial software package used for Kentucky school districts), Jefferson County's testing budget for 2001-2002, and survey responses from Program Review staff's surveys of teachers, principals, and superintendents.⁴

Local School Districts' Costs

The estimate for the cost of the assessment for school districts includes administrative and personnel expenditures.

Staff used two main components to calculate the cost of the CATS assessment to local school districts: personnel costs and administrative costs. Most school districts utilize the object code "0646 Tests" under the MUNIS system to report administrative expenditures associated with CATS. Because the MUNIS category is not specifically designated for CATS tests, it is possible that some administrative costs assigned to this category are for other types of assessments, such as English as a Second Language. Most of the expenses in this category are likely to be associated with the CATS assessment, however. A few school districts did not list any expenses under this category. According to a KDE representative, these districts' CATS-related administrative expenses could have been listed under broader categories such as "0640 Books and Periodicals" or "0610 Supplies." Jefferson County does not use the MUNIS reporting system, so its testing expenditures were obtained from a representative of the school district. Total district administrative expenses for the CATS assessment were calculated from those totals in category "0646 Tests" and the totals received from Jefferson County.

Personnel costs include replacing teachers who attended portfolio training or were portfolio scorers and portions of the superintendent's and district assessment coordinator's salary.

Three categories of personnel expenses were included in the estimate of district cost:

- The cost of hiring replacements for teachers who were attending portfolio scoring training or who were scoring the portfolios,
- The portion of the district assessment coordinator's salary associated with assessment duties, and
- The share of the superintendent's salary associated with assessment duties.

⁴ A 2002 survey administered by the Office of Education Accountability asked superintendents about the local cost of student assessment. Program Review staff were unable to use this survey to determine the cost for local districts because there was no standard for what should be considered a cost of implementing the CATS test.

Personnel cost estimates were based, in part, on responses received on the Program Review surveys of principals and superintendents.

Because records were not available for this information, it was necessary to make estimates based on results from the surveys of principals and superintendents.

Relevant Survey Questions for Principals

- Do you hire substitute teachers to replace teachers who are attending portfolio-scoring training? If yes, how many and for how many days?
- Do you hire substitute teachers to replace teachers who are scoring writing portfolios? If yes, how many and for how many days?

Relevant Survey Questions for Superintendents

- In terms of the direct costs of CATS testing and assessment to your school district, what percentage of your time is devoted to specific CATS testing and assessment tasks?
- Does your district have a District Assessment Coordinator? If yes, what percentage of his/her time is devoted to specific CATS testing and assessment tasks?

The answers of superintendents and principals to these questions were used to develop the following assumptions for the cost estimate:

- Thirty-eight percent of Kentucky schools hire substitutes for teachers undergoing writing portfolio training. The average was 3.5 substitutes for 1.4 days of instruction. This equates to about five days of instructional expenses for training for scoring of writing portfolios.⁵
- Fifty-six percent of schools arrange for substitutes for days when teachers are scoring portfolios. The average was 5.1 substitutes for 1.8 days of instruction. This represents about nine days of instructional expense for writing portfolio scoring.

The cost of hiring substitutes was based on the average daily salary of \$198 for teachers in 2001-2002.⁶ To estimate the district personnel cost of writing portfolio training, the average daily salary of \$198 was multiplied by 5 days for 491 schools (38 percent of 1,292 total schools). To estimate the district cost of

⁵ It should be noted that the instructional expenses in days for training and portfolio scoring are averages per school for substitutes. Principals and superintendents were not asked how many teachers were substituted for—or over how long a period.

⁶ The daily salary was based on the average annual teacher's salary of \$37,105 for an average of 187.4 contract days for school year 2001-2002. It would have been more accurate to use the average daily salary—presumably lower—of substitute teachers, but this data was unavailable.

writing portfolio scoring, \$198 was multiplied by 9 days for 723 schools (56 percent of 1,292 schools).

Based on responses from the survey, superintendents spend approximately 28 percent of their time on specific CATS testing and assessment tasks and district assessment coordinators spend approximately 44 percent of their time on such tasks. The statewide totals for superintendents' and district assessment coordinators' salaries for the 2001-2002 school year was provided by KDE from the MUNIS financial reporting system. Twenty-eight percent of the total for superintendents' salaries and 44 percent of the total for district assessment coordinators' salaries were then calculated.

The components of the local cost estimate are shown in Table 2.2. The cost of the CATS assessment for local school districts in 2001-2002 was estimated to be \$10.6 million—about \$16 per enrolled pupil, or \$22 per student tested.⁷

Table 2.2
Estimate of School Districts' CATS Assessment Costs,
2001-2002

Expenditure Category	Estimated Cost (in \$million)
MUNIS "0646 Tests" and Jefferson County	3.9
Salaries for Substitute Teachers	1.8
28% of Total Superintendents' Salary	4.5
44% of Total District Assessment Coordinators' Salary	0.4
Total	\$10.6 million

Source: Calculated by Program Review staff based on the survey of superintendents and data from KDE, MUNIS, and Jefferson County.

The \$10.6 million figure should be considered a rough estimate of costs borne by local school districts for CATS assessments. The MUNIS category that was used contains expenditures for other kinds of tests. The number of schools using substitutes to cover for teachers doing writing portfolio training or scoring is based on surveys from samples of school administrators. The percentage of time spent on CATS assessment by superintendents and district assessment coordinators is also based on a survey of a sample of administrators. Samples have error rates by their nature.

⁷ Expenditure per pupil is calculated by dividing total local expenditures by 654,363—the number of K-12 public-school students in the state in 2001-2002. Expenditure per student tested is calculated by dividing expenditures by 474,182—the number of students who took CATS tests in 2001-2002.

Even if all the potential errors in the procedures used here to calculate local costs canceled each other out and \$10.6 million really is the true local cost of CATS assessment, this report's estimate is a one-time affair. There is no reason to assume that costs do not change over time. It is also hard to argue that local districts' costs and any trend in those costs would not be useful information for legislators, district officials, and the public. Therefore, this report recommends that KDE develop procedures to gauge districts' annual assessment expenditures and publish this information.

Recommendation 2.4

KDE should create a CATS testing expenditure category in the MUNIS system and encourage school districts to utilize this category for all CATS administrative expenses. Any problems in implementing this change should be reported to the Education Committee and the Education Assessment and Accountability Review Subcommittee.

State Costs

As shown in Table 2.3, state government's total cost for CATS assessment was about \$10.3 million—\$15.75 per enrolled pupil or \$21.75 per student tested—in fiscal year 2003.⁸ State costs may be estimated more precisely because one state agency—KDE's Office of Assessment and Accountability—has primary responsibility for Kentucky's testing system. There are two components to the General Fund cost estimate for the state: personnel and nonpersonnel expenditures. Included in these categories are funds for contracts for the following:

- CTB McGraw-Hill scoring and auditing;
- HumRRO, an independent research firm hired by KDE to help evaluate the CATS assessment; and
- The University of Kentucky Alternate Portfolio Master Agreement (M-02157539).

The estimated local cost was for the 2001-2002 school year and the state cost is for the 2002-2003 fiscal year, so the time periods are

⁸ The number of students tested in the 2002-2003 school year was unavailable at the time of this report's writing. The number of tested students in the 2001-2002 school year was used to calculate the state cost per tested student in fiscal year 2002-2003.

different. Given that caveat, adding the local and state figures, it is estimated that the total cost per year of CATS assessment is about \$20.9 million, or about \$32 per pupil. No other study has tried to estimate the total cost of the CATS assessment, although a National Bureau of Economic Research study (Hoxby, 2002) estimated the state cost to be about \$18 per pupil, slightly higher than the estimate provided in this report.

Table 2.3
Estimate of State General Fund Expenditures for CATS Assessment, FY 2003

Expenditure Category	Cost
Office of Assessment and Accountability: Personnel	\$9,500,200
Office of Assessment and Accountability: Other	\$821,167
Total	\$10,321,367

These figures include the following contracts: CTB McGraw-Hill (\$7,042,225), HumRRO (\$239,000), and UK Master Agreement (\$250,000).

Source: Compiled by Program Review staff from data supplied by KDE.

Other States' Assessment and Accountability Systems

Education Week researchers survey state education administrators to produce *Quality Counts*, a comprehensive annual report of the states.

This chapter has described Kentucky's accountability and assessment system. Obviously, it is not the only such system. The remainder of this chapter provides a brief overview of others states and how Kentucky compares.

Education Week researchers survey state education administrators to produce *Quality Counts*, a comprehensive annual report of the states. *Quality Counts 2003* includes each state's grade for standards and accountability. Forty percent of the grade is based on the state's standards. The remaining 60 percent of the grade is based on assessment and accountability.

Among the states, *Quality Counts 2003* ranked Kentucky third in the nation for its standards and accountability system.

Kentucky's overall grade on standards and accountability was a 94, ranking Kentucky third, with only Florida and New York receiving higher grades. To grade states on assessment, *Quality Counts 2003* measures:

- The types of test items the state used in 2002-2003 to measure student performance;
- The subjects in which the state used criterion-referenced assessments aligned to state standards in 2002-2003;
- Whether state criterion-referenced assessments underwent external alignment review—an independent review to determine the degree to which test content matches subject area content—in 2002; and

- Whether the state participated in the 2002 National Assessment of Educational Progress.

State Testing

Types of Tests. *Quality Counts 2003* graded states on their use of criterion and norm-referenced tests. Criterion-referenced tests assess how well students know certain content; norm-referenced tests are used to rank students. States have increased their use of criterion-referenced tests. For example, *Quality Counts 2003* reports that the number of states qualifying as administering criterion-referenced tests in English and mathematics at the elementary, middle, and high school levels increased from 37 states in 2002 to 42 states in 2003.

Thirty states, including Kentucky, use a combination of norm- and criterion-referenced tests.

Thirty states, including Kentucky, use norm-referenced and criterion-referenced tests. Two states rely solely on norm-referenced tests for assessment and 18 states administer only criterion-referenced tests.

Types of Questions Used in State Tests. Norm-referenced tests are typically multiple choice. Most criterion-referenced tests include a multiple-choice component, but also include item types such as short answer, extended response, and portfolios. A short-answer question asks a student to answer a question with a word or two. An extended-response question asks a student to exhibit knowledge through more writing. Table 2.4 summarizes the types of test items states use. Every state except Iowa, North Carolina, and Montana uses two or more types of test items. Kentucky uses multiple choice, extended response—referred to as open response—and writing portfolios on the CATS test. Forty-four states use multiple-choice and extended-response questions, and 32 state tests contain short-answer items. Kentucky and Vermont are the only states that include writing portfolios in their state assessments.

Table 2.4
Types of Test Items Used in States' Assessment Systems

Item Type	Number of States	Kentucky?
Multiple Choice	49	Yes
Extended Response	45	Yes
Short Answer	32	No
Portfolio	2	Yes

Source: Compiled by Program Review Staff from *Quality Counts 2003*.

The four most common subjects tested with criterion-referenced tests are language arts, mathematics, science, and social studies and history.

Subjects Tested. Four subjects are most often tested with criterion-referenced tests. A total of 48 states test students in language arts, 46 states test students in mathematics, 34 states test students in science, and 24 states test students in social studies or history. Kentucky tests students in all four subjects measured by *Quality Counts*.

Accountability Systems

To grade states on accountability, *Quality Counts 2003* measures the type of information used to evaluate schools and whether:

- The state requires school-level report cards to report results by race, poverty, limited English proficiency, and disability;
- The high school report card shows graduation rates for these groups;
- The state has a statewide student identification system;
- The state assigns ratings to all schools or identifies low-performing schools;
- The state provides assistance to low-performing schools;
- The state provides rewards to high-performing or improved schools; and
- State accountability includes sanctions such as school closure, reconstitution, reconstituting schools as charters, permitting student transfers, turning over schools to private management, and withholding funds.

As of 2003, 47 states required school-level report cards. As shown in Table 2.5, about half of these states do not report student performance data by race, poverty level, limited English proficiency, or disability. Kentucky does so for all four groups.

Table 2.5
State Report Cards That Report Student Performance Data by Groups

Student Performance Data	Number of States	Kentucky?
Disability	29	Yes
Race	27	Yes
Poverty	24	Yes
Limited English Proficiency	24	Yes

Source: Compiled by Program Review staff from *Quality Counts 2003*.

Program Review and Investigations

Twenty-nine states assign ratings to all schools or identify low-performing schools.

Twenty-nine states assign ratings to all schools or identify low-performing schools. As Table 2.6 shows, 17 states use test scores and other information to evaluate schools. Like Kentucky, most of these states use test scores and nonacademic factors such as attendance, retention, and dropout rates to evaluate school performance.

Table 2.6
Information Used by States To Evaluate Schools

Type of Information	Number of States	Kentucky?
Test Scores and Other Information	17	Yes
Test Scores of Specific Subgroups	9	Yes
Site Visits or Reviews	3	No

Source: Compiled by Program Review Staff from *Quality Counts 2003*.

Twenty-two state accountability systems include sanctions to schools based on performance and 17 states provide schools with rewards based upon performance. Twenty-seven states, including Kentucky, provide assistance to low-performing schools.

Chapter 3

Evaluating CATS

A full-fledged evaluation of the Commonwealth Accountability Testing System (CATS) is beyond the scope of this study, but some research questions approved by the Program Review and Investigations Committee are evaluative. Specifically, some questions on the surveys of Kentucky educators call for judgments of various parts of CATS. This chapter reviews several such questions from the surveys. Committee members also directed staff to compare scores on the CATS tests to other tests of Kentucky students. A comparison of results from CATS and ACT tests is provided. Committee members also asked whether the National Technical Advisory Panel on Accountability and Assessment had judged CATS valid. In a document prepared for this report, panel members agreed there is substantial evidence that the Kentucky Core Content Test and CATS are valid and reliable as used.

The Surveys of Educators

The Program Review and Investigations Committee instructed staff to survey Kentucky teachers, principals, and superintendents to elicit their opinions about various aspects of CATS.

The committee instructed staff to survey Kentucky educators to elicit their opinions about various aspects of CATS. About 800 teachers (38 percent of those who were sent questionnaires), 500 principals (42 percent), and 100 superintendents (72 percent) completed questionnaires. Details on the surveys are provided in Appendix D. This chapter reviews responses to questions on the following topics:

- How CATS affects the quality of education, the curriculum, the way teachers teach, and how students learn;
- Whether question repetition affects how teachers prepare students for the test;
- The process used to set school improvement goals;
- The consequences to schools that fail to improve and the quality of assistance provided to “needs assistance” schools by KDE; and
- The academic and nonacademic components included in the accountability index.

The Effects of CATS on the Quality of Education

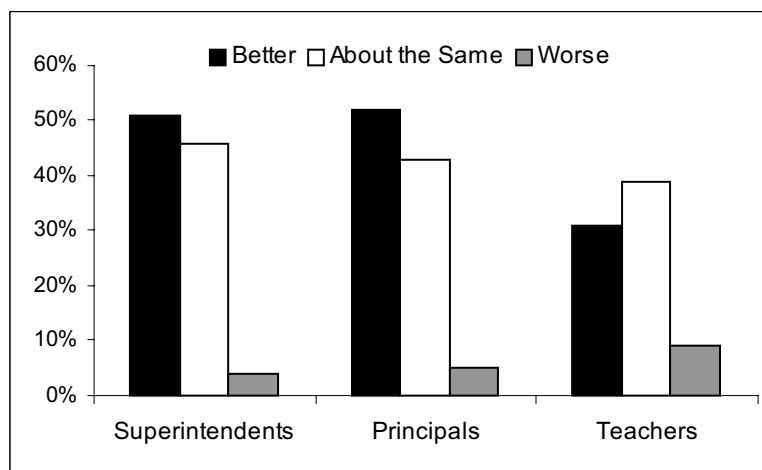
The surveys included several questions that called for educators to express judgments on the effects of CATS on the quality of education in general and on aspects of education such as teaching, curriculum, and preparing students for tests. Majorities of teachers, principals, and superintendents answered that the quality of education is the same or better under CATS compared to the previous accountability system. Majorities of principals and superintendents reported that CATS affects teaching and learning positively and that they were satisfied with the Core Content for Assessment. Teachers were less supportive than principals and superintendents of the effect of CATS on teaching, learning, and the core content, but their most common responses to these questions were still positive.

More than half of principals and superintendents responded that education is now better than under the previous system. About a third of teachers said that education is better under CATS.

Comparison With the Previous Accountability System.

Teachers, principals, and superintendents were asked their views on how CATS has affected the overall quality of education compared to the Kentucky Instructional Results System, the accountability system used before CATS. As shown in Figure 3.A, more than 50 percent of principals and superintendents responded that the quality of education is now better. More than 40 percent of each group said that the quality of education is the same as before. About a third of teachers said that education is better under CATS; 39 percent said it is about the same as before.

Figure 3.A
Compared to the Previous System, the Quality of Education Is Now Better, About the Same, or Worse?



Source: Program Review staff surveys of teachers, principals, and superintendents.

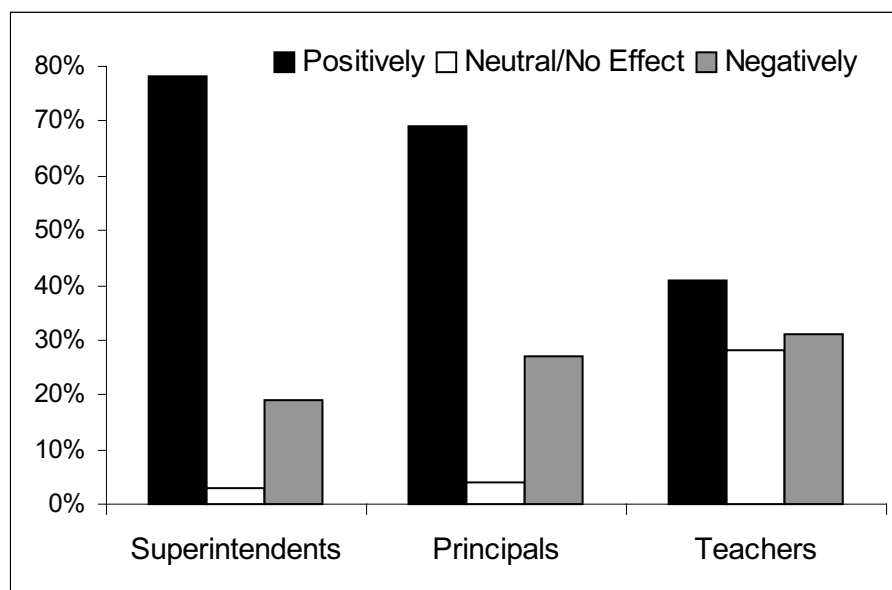
More than 75 percent of principals and superintendents reported that CATS has positively affected curriculum.

Effect on Curriculum. Kentucky principals and superintendents were asked how CATS affects school and district curricula. Large majorities, 76 percent of principals and 79 percent of superintendents, indicated that CATS has somewhat positively or very positively affected curriculum, or what is actually taught in schools. Eighteen percent of principals and 16 percent of superintendents answered that CATS has had a negative impact on curriculum.

Most principals and superintendents indicated that the CATS test's effect on teaching was positive. Less than half the teachers agreed.

Test's Effect on Teaching. Educators gave divergent answers when asked how the CATS test affects the way teachers teach. Superintendents evaluated the CATS test most positively; teachers were the most negative. The responses of superintendents, principals, and teachers are summarized in Figure 3.B. More than two-thirds of principals (69 percent) and superintendents (78 percent) responded that the CATS test affects teaching somewhat or very positively. Only 41 percent of teachers agreed that the effect of the CATS test on teaching was positive. Almost a third of teachers said that the impact is negative, compared to 27 percent of principals and 16 percent of superintendents.

Figure 3.B
How Does the CATS Test Affect the Way Teachers Teach?



Source: Program Review staff surveys of teachers, principals, and superintendents.

The majority of teachers, principals, and superintendents responded that CATS positively affects what students learn.

Effect on Student Learning. The majority of teachers, principals, and superintendents responded that CATS positively affects what students learn. Teachers were again less positive than the other educators. About three-fourths of principals and superintendents responded that CATS somewhat or very positively affected what students learn. Around 20 percent of each group responded that the effect was somewhat or very negative. A little more than half the teachers—55 percent—indicated the effect was positive and more than one-fourth said the effect was negative.

About half of the teachers indicated they were at least somewhat satisfied with the core content.

Core Content for Assessment. The Core Content for Assessment details the material on which students will be tested. When asked about the core content, slightly more than half of the teachers indicated they were at least somewhat satisfied with it. A little more than one-fourth was dissatisfied. Large majorities of principals (73 percent) and superintendents (84 percent) indicated they were satisfied with the core content.

More than 60 percent of teachers, principals, and superintendents reported that teachers do not have adequate instructional time to teach the core content.

Teachers, principals, and superintendents were more in agreement when asked whether teachers have enough instructional time to adequately teach the core content. More than 60 percent of each group said that teachers do not have adequate instructional time to teach the core content.

Repetition of Test Questions. Approximately 80 percent of the Kentucky Core Content Test's questions are repeated from one year to the next. Questions on the Comprehensive Test of Basic Skills norm-referenced test have not changed since 1997, even though alternate forms of the test are available. The evidence from responses to Program Review staff's survey of educators suggests that repetition of questions from year to year does affect teaching for some teachers. It is unlikely that teachers would volunteer on a questionnaire that they prepare students with questions from past tests. Given that, there is little evidence from the survey to indicate that teachers are "teaching the test" if that is defined as teachers incorporating unreleased past questions into their lessons. Based on the volunteered follow-up responses of teachers to the question, it is possible that many answers were not based on the repetition aspect of the survey question anyway.

Program Review staff asked principals, superintendents, and teachers—with a different version of the question for teachers—if the repetition of CATS questions affects the way teachers prepare students for the test. Principals and superintendents were asked if

preparation for the test was affected by repetition of questions, and if yes, how many teachers were affected—a few or many. For each group, more said preparation was affected than said it was not affected. Just fewer than half the principals responded that repetition affected preparation for a few or many teachers. Fifty-seven percent of superintendents reported that question repetition affected preparation for a few or many teachers.

More teachers, principals, and superintendents said student preparation for the CATS test was affected by the repetition of questions than said it was not affected.

Each teacher surveyed was asked how the repetition of questions affects the way “you” prepare “your” students for the test. Just fewer than half the teachers said question repetition did not affect preparation. Only 7 percent said that test preparation was affected a lot by having questions repeated; 38 percent said that preparation was affected somewhat.

Space was provided for teachers to explain their answer to this question and more than 300 (out of 800 respondents) did so. Most of those who provided explanations said question repetition does not affect preparation of students. Many gave answers that explicitly rejected the use of test questions, such as “No, because we are not allowed to teach to the test.” Others explained that their teaching was centered on the core content.

Many teachers did answer that repetition affects how they prepare students. Only a handful gave answers that could be interpreted as meaning they use repeated questions to prepare students for specific questions, however. Most of those who responded that repetition affects preparation noted that question repetition helped them better prepare students for the format of questions or for content that is stressed on the test. It is important to note that teachers could gather these types of information from past tests whether questions are repeated or not. Given such responses, it is unclear how many teachers, principals, or superintendents focused on the repetition aspect of this question when answering.¹

KDE on Why the Repetition of CATS Questions Is Necessary

Kentucky Department of Education representatives told Program Review staff that the repetition of questions is necessary to equate student performance. To measure how well students perform on the test in one year compared to the previous year, at least one complete test form and parts of two others must stay the same—using the same questions—to make the comparison valid.

¹ There was no provision on the surveys of principals and superintendents for follow-up responses to this question.

KDE staff also indicated that it is very difficult to replace items with new ones because all new items must be field tested to ensure reliability and validity. KDE field tests approximately 24 questions per content area per year. Questions that have been field tested as reliable and valid are put into an item pool for consideration to replace some questions that have been repeated.

The use of a large item pool and multiple forms is necessary to adequately cover the breadth of the core content. KDE staff indicated that the large item pool produced by the use of six forms of the Kentucky Core Content Test and the heavy weight put on open-response questions would make it very difficult for teachers to teach students the test. Teachers do not know the form of the test each student will take.

KDE staff indicated that it would be easier to cheat on multiple-choice questions than open-response questions. For teachers to show students how to answer an open-response question, it would be necessary for students to learn a core concept. For example, if a student was asked to answer an open-response science question explaining the components of an ecosystem, the teacher would have to teach that student the components and relationships of an ecosystem.

Finally, KDE representatives told Program Review staff that the differences over time in scores for repeated items are very small. KDE staff expressed confidence that when cheating does occur, an administrator, teacher, or parent will find out and contact KDE.

Setting Improvement Goals

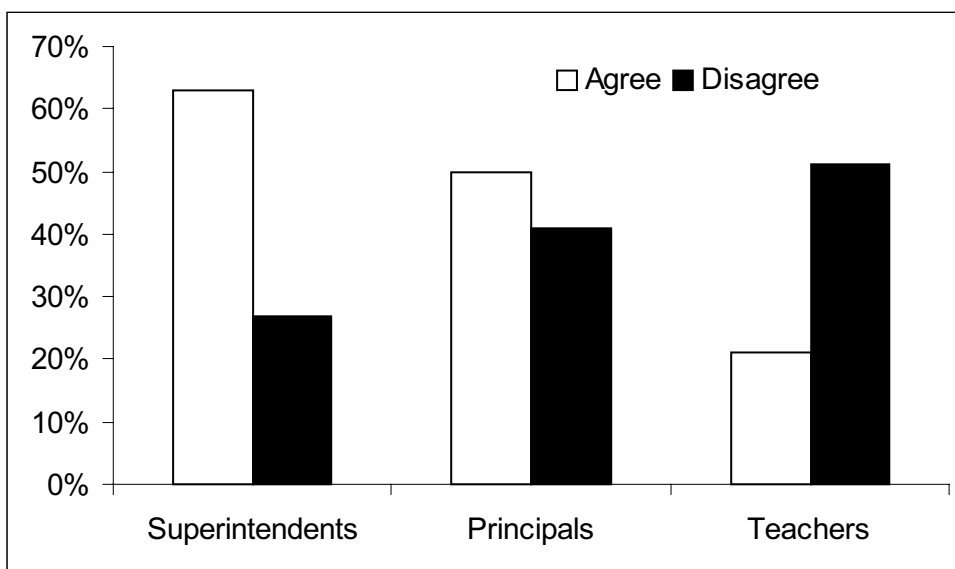
More than half of teachers disagreed that the process used to set school improvement goals was appropriate. Based on their follow-up answers, many teachers appeared to be commenting on the appropriateness of the goals, not just the process used to create them.

Teachers, principals, and superintendents were asked whether the process used by the Kentucky Board of Education to set school improvement goals was appropriate. Teachers were asked a follow-up question requesting that they explain their answer. A review of their explanations shows that respondents seemed to be commenting on the appropriateness of the improvement goals, not just the process used to create them. The description that follows is based on the wording of the question on the survey. It should be kept in mind, however, that educators are apparently giving their views on the process, the goals, or both.

Figure 3.C summarizes the responses to this question. Teachers voiced the strongest disagreement with the process. More than half the teachers disagreed that the process was appropriate, including

19 percent of respondents who strongly disagreed. About one-fifth of teachers agreed that the process was appropriate and only 2 percent—one teacher in 50—strongly agreed with the statement.

Figure 3.C
Do You Agree or Disagree That the Process Used by the Kentucky Board of Education To Set Improvement Goals Was Appropriate?



Source: Program Review staff surveys of teachers, principals, and superintendents.

Principals were more evenly divided on this issue. Half agreed that the process was appropriate, but 41 percent disagreed. Superintendents were the most positive. More than 60 percent agreed that the process to set improvement goals was appropriate; 27 percent disagreed.

Many teachers responded that improvement goals are unrealistic because the system does not measure individual students' progress.

When asked to explain their answer, almost 400 teachers did so. More than half of those answering gave a response that was classified in one of these three categories:

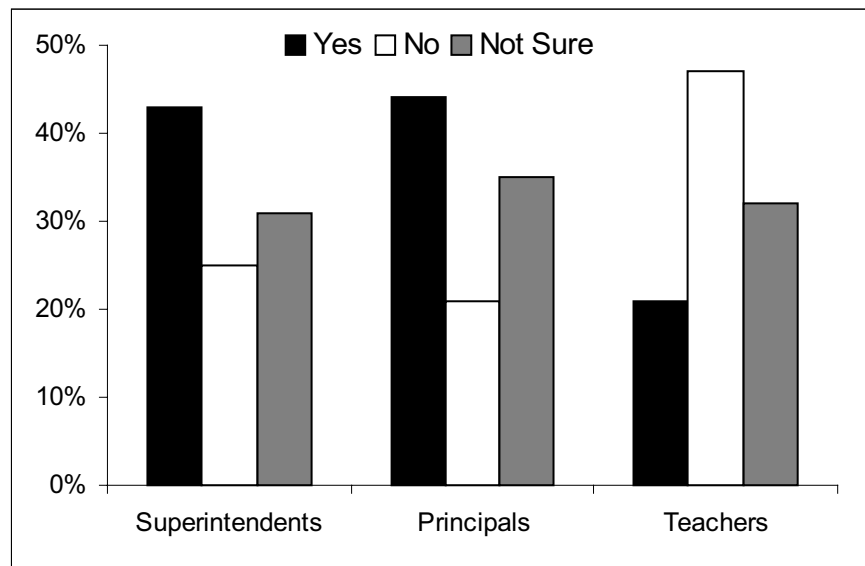
- It is impossible to reach the proficiency goal because not all students are capable enough;
- The test does not measure progress of the same students over time; and
- Student population characteristics are not reflected in goals.

About 20 percent of teachers and more than 40 percent each of principals and superintendents reported that their schools could reach proficiency by 2014.

As predicted by these comments, teachers were the least optimistic when asked whether their school could reach proficiency by 2014. As indicated in Figure 3.D, only 21 percent of teachers responded that their school could reach the goal. More than 40 percent each of superintendents and principals answered that proficiency by 2014 was attainable for their school (principals) or all schools in

the district (superintendents). Almost half—47 percent—of teachers reported that their school could not attain proficiency by the goal year. More so than for almost any other question on the surveys, many educators responded “not sure” to this question. About a third each of superintendents, principals, and teachers indicated that they did not know whether schools would attain proficiency by 2014.

Figure 3.D
Can Your School Reach Proficiency by 2014?*



*The question for superintendents was whether all schools in the district could achieve proficiency by 2014.

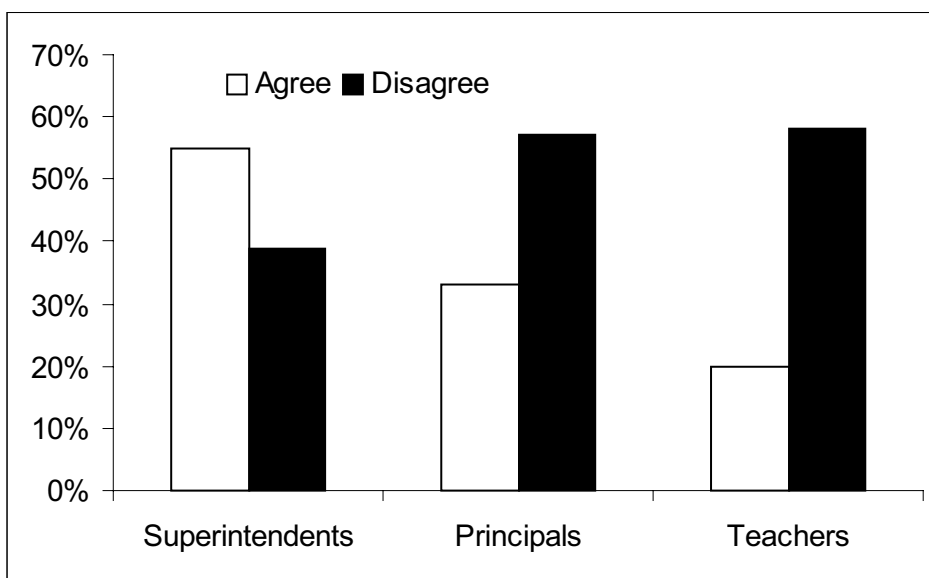
Source: Program Review staff surveys of teachers, principals, and superintendents.

Schools That Do Not Meet Improvement Goals

More than 55 percent of teachers and principals disagreed with the statement that the consequences to schools that fail to meet their biennial improvement goals are appropriate.

Consequences. Educators were asked to agree or disagree that the consequences to schools that fail to meet their biennial improvement goals are appropriate. Their answers to the question are summarized in Figure 3.E. More than 55 percent each of teachers and principals disagreed with the statement that the consequences are appropriate, with about 20 percent of each group strongly disagreeing. Only one-third of principals and one-fifth of teachers agreed that consequences are appropriate. Superintendents were more supportive of the consequences. A majority of superintendents, 55 percent, responded that the consequences to schools were appropriate. About 40 percent disagreed.

Figure 3.E
Do You Agree or Disagree That the Consequences to Schools That Fail To Improve Are Appropriate?



Source: Program Review staff surveys of teachers, principals, and superintendents.

The majority of teachers and principals who have worked in a school classified as needing assistance indicated that the aid provided by their school districts and KDE was helpful.

Quality of Assistance. KDE and school districts provide help to schools that are classified as needing assistance. Principals and teachers whose schools are receiving or have received assistance were asked to rate the helpfulness of this aid. The majority of teachers and principals who have worked in schools classified as needing assistance indicated that the assistance provided was somewhat or very helpful. Principals responded more positively, with 75 percent rating the assistance as helpful, compared to 63 percent of teachers.

The Accountability Index

Program Review staff elicited opinions from teachers, principals, and superintendents about the appropriateness of the weights assigned to academic and nonacademic components of the CATS accountability index.² For each component, educators were asked whether the weight was “about right,” “too high,” or “too low.” The weight for each component was not provided on the survey form because the weighting of components differs by level of school.

² Table 2.1 on page 17 of this report shows how each component is weighted by level of school.

Responses to the questions are summarized in Table 3.1. The table shows the percentages of respondents who said the weight of each item was too low or too high. If the total percentages shown in the table for a component are low, then a large percentage of educators said the weight for that item are about right.

Table 3.1
Educators' Opinions of the Weighting of Accountability Index Items
(as Percentage of Members of Each Group Answering Each Question)

Item	Teachers		Principals		Superintendents	
	Too Low	Too High	Too Low	Too High	Too Low	Too High
Reading	4%	9%	5%	7%	13%	7%
Math	5	10	4	9	14	7
Science	5	12	2	13	7	17
Social Studies	4	11	2	12	8	17
Writing Portfolios	16	48	26	31	13	38
On-demand Writing	6	37	9	35	13	25
Arts and Humanities	8	31	7	35	7	42
Practical Living/Vocational	10	25	6	33	7	38
Attendance	8	26	9	20	13	15
Retention Rate	9	33	5	29	8	23
Dropout Rate	5	38	4	24	11	25
Transition to Adult Life	12	24	5	23	6	30
CTBS Testing	18	8	38	3	45	5

Source: Program Review staff surveys of teachers, principals, and superintendents.

There is less disagreement among superintendents, principals, and teachers on the weighting of items in the accountability index than on many other issues about which they were asked. Based on their responses, most of the items can be separated into three distinct groups:

- Items for which there is widespread agreement that the weights are appropriate,
- Items that a relatively high percentage of educators report are weighted too highly, and
- One item that a relatively high percentage of educators indicate is not weighted highly enough.

At least 75 percent each of teachers, principals, and superintendents responded that the weights for reading, math, science, and social studies are about right.

At least 75 percent of each group responded that the weights for reading, math, science, and social studies are about right. Most of the teachers and principals who did not agree that the weights were about right for these items thought the weights were too high.

According to at least 20 percent of each group, seven items are given too much weight on the accountability index: writing portfolios, on-demand writing, arts and humanities, practical living/vocational skills, the retention rate, the dropout rate, and transition to adult life. At least one-fifth of teachers and principals wrote that the weight for attendance is too high. Almost half of teachers said writing portfolios are weighted too highly—the largest percentage of any group saying an item is weighted too highly. At least 30 percent each of principals and superintendents agreed. It should be noted, however, that more than 10 percent of teachers and superintendents and 26 percent of principals answered that the weighting of writing portfolios is too low. At least 30 percent of each group—42 percent of superintendents—said that the arts and humanities component is too highly weighted.

More than one-third of principals and superintendents, 38 and 45 percent, respectively, answered that CTBS norm-referenced testing is weighted too low. Eighteen percent of teachers agreed. For each group of educators, this was the item that the largest percentage thought is not weighted highly enough.

When asked to comment on which components should be removed from the accountability index, more than half of the 317 teachers who answered indicated the writing portfolio. Some said the portfolio should be removed altogether; some indicated it should be removed only on the fourth-grade test. More than one-fifth of superintendents and more than one-third of principals who listed at least one component to be removed mentioned portfolios, too. There were no other index components for which superintendents, principals, and teachers were in such relative agreement.

Student Accountability

Responding to a question asking what they would like to change about CATS and in their follow-up comments for numerous questions, many educators expressed the need for student accountability on the CATS test. Comments from teachers, principals, and superintendents such as the ones below were typical.

**Comments From the CATS Surveys:
Changes That Should Be Made to CATS**

“Put more pressure on the students to perform and less on the teachers.”

“Find a way to make students accountable for how well or poorly they do on the test. Right now I have students, great students, who don’t care about the test and don’t do their best because it has NO impact on them one way or another.”

“Find a way to build some student accountability into the process (e.g. utilize CATS scores in the KEES program).”

KDE has proposed including KCCT and writing-portfolio scores in high school students’ grade point averages.

Students are not held directly accountable for test scores. According to the CATS 2002 Interpretive Guide, KDE has considered a number of proposals to increase student accountability. At the June 2003 meeting of the National Technical Advisory Panel on Assessment and Accountability (NTAPAA), KDE officials discussed with NTAPAA members a proposal intended to increase student motivation on the CATS test. School districts would have the option of having scores from the Kentucky Core Content Test (KCCT) and the writing portfolio count on a student’s high school grade point average (GPA).

Scores from the KCCT would be a factor in the overall GPA of high school students. The performance levels from the KCCT used in this calculation would be those assigned by the scoring contractor. Individual student performance on the KCCT and writing portfolio could count for no more than 3 percent of the maximum GPA. Also, performance could only add to the GPA; a bad performance would not reduce GPA.

NTAPAA members expressed concern that KCCT and writing-portfolio scores would only add to a student’s GPA. If a student scored novice on all academic components of KCCT, his or her GPA would remain the same. The members felt that integrating CATS scores into students’ GPAs could possibly benefit high-achieving students who wanted to further increase their GPAs, but have little effect on the lower-achieving students.

Comparing CATS and ACT Scores

Comparing the scores of students on the CATS assessment to other tests is a way to address the question of whether CATS is a valid measure of a student's academic performance. Of the other tests, it is only feasible to compare CATS and ACT scores at the student level.

Comparing the scores of students on the CATS assessment to other tests is a way to address the question of whether CATS is a valid measure of a student's academic performance. National tests of academic aptitude for potential comparisons include the ACT, the Scholastic Aptitude Test (SAT), the National Assessment of Educational Progress (NAEP), and the Armed Services Vocational Aptitude Battery (ASVAB). Only the results from the ACT are compared to CATS scores in this report due to limitations in the samples used and results available from the other tests.

High school students who plan to attend a four-year college or university typically take the ACT or SAT. The explicit purpose of these tests is to predict success in a college or university. In Kentucky, public universities use ACT scores much more often than SAT scores for admission decisions; therefore, the SAT is taken by too small a group of Kentucky students for use as a comparison to CATS scores. The NAEP test is administered to only a sample of Kentucky students, typically fewer than 3,000 students statewide. In addition, separate samples are used for each of the NAEP academic components, such as mathematics, reading and writing. Because the NAEP is taken by a sample of Kentucky students and the students vary by subject area, NAEP scores and CATS scores can be compared at the state level, but not at the student level. The Department of Defense administers the ASVAB to all people wishing to enlist in the Armed Services. A number of Kentucky students take this test each year, but the Department of Defense does not release individual results.

It is feasible to compare the CATS and ACT scores of individual students. The format and content of the ACT and CATS assessments differ considerably, however.

It is feasible to compare the CATS and ACT scores of individual students. There are caveats that should be kept in mind, however. First, almost all public school students in Kentucky take the CATS test. Typically, students taking the ACT are those interested in attending a four-year college or university. Second, the CATS assessment is designed to measure a school's achievement based on students' scores. The ACT is designed to predict the success of college for individual students and school-level indexes are not computed. Third, the CATS assessment is not reflected in the student's GPA or used for college admission, so it is likely that a student's motivation will differ compared to the ACT assessment, which affects a student's admission prospects to chosen schools. The tests themselves are different as well. The ACT is comprised of equally weighted multiple-choice questions. The CATS assessment is made up of multiple-choice and open-response questions, writing portfolios, and on-demand writing. The ACT

has four subtests: reading, mathematics, science reasoning, and English. The CATS assessment includes the academic components described earlier in this chapter.

Despite these differences, the ACT and CATS assessments are designed to measure a student's ability to apply academic content knowledge. It is unclear to what degree results from the two tests should be related, but there is no obvious reason why the scores should not be correlated.

Trends in Statewide ACT and CATS Scores

To establish an appropriate comparison of ACT and CATS scores over time, precautions must be taken so that the comparisons are valid. Only public school students participate in the CATS assessment, however, any public, private, or out-of-state student may take the ACT in Kentucky. CATS scores include nonacademic factors such as dropout rates and CTBS scores.

ACT and CATS scores from students who took the ACT from 1999 to 2002 were used to measure trends in scores.

To avoid these comparison problems, only CATS scores, less the nonacademic factors and the CTBS scores, from public school students who took the ACT in 1999 to 2002 were included in this calculation of state average scores. As shown in Table 3.2, statewide average CATS scores for these students have increased over the past four years, up 50 points to an index value of 400 in 2002. Over the same time period, statewide ACT scores have remained virtually unchanged.

Table 3.2
Kentucky Statewide Average ACT and
CATS Scores for Public School Students
Taking the ACT, 1999 to 2002

Year	Number of Students	Average ACT Score	Average CATS Score
1999	22,528	20.0	350.9
2000	24,023	20.0	390.6
2001	23,922	20.0	397.4
2002	24,182	19.9	400.8

Source: Program Review staff analysis of data provided by KDE.

The Correlation Between Students' ACT and CATS Scores

Student-level CATS scores were calculated for students who had participated in three years of testing.

Making a valid comparison between CATS and ACT scores is complicated. The details of the work necessary to do so are contained in Appendix E. CATS scores are calculated at the school level; ACT scores are student-level scores. To see if students' scores on the two tests are correlated, student-level CATS scores must be calculated. The formula for the school-level calculation was used. Eight components remained after the subtraction of nonacademic factors and CTBS scores that are not relevant to students' individual scores. For a student to have a score for each of these eight components, he or she must have been tested in the 10th, 11th, and 12th grades. CATS scores are available for 1999 through 2002, therefore only those students who began 10th grade in 1999 or 2000 would have scores for all three years of testing. These student-level scores are matched to individual ACT scores. The resulting data set consists of more than 41,000 students who took the ACT from 1999 to 2002 and who had scores from the necessary eight components from the CATS test. It should be noted that this group of students differs from the student population as a whole. For example, for those not taking the ACT, the average CATS composite score was 365.5. Those taking the ACT had a CATS score of 398. Females made up 58 percent of those taking the ACT, compared to 44 percent of those who did not take the ACT.

The correlation between individual students' ACT and CATS scores is relatively high.

Table 3.3 shows the correlation coefficients for a CATS composite score calculated based on scores on the eight KCCT components and ACT composite scores.³ Coefficients are also shown for the math and reading components of the tests. With a correlation coefficient of .738, the composite CATS and ACT scores are relatively highly correlated. Of the math and reading components, correlation is higher for math (.696) than for reading (.550).

³ A correlation coefficient, which can range from -1 to +1, measures the extent to which the values of two or more measurements go up and down together. Perfect negative correlation (-1) means that scores on one measure are the opposite of another. For example, high scores for one measure are associated with low scores on another. Perfect positive correlation (+1) means that the two measures are comparable in a relative sense. For instance, high scores for one measurement are associated with high scores on another. A coefficient of 0 means there is no correlation between the measures.

Table 3.3
Correlations Between CATS and ACT Scores:
Composite, Math, and Reading

	ACT Composite	ACT Math	ACT Reading
CATS Composite	.738		
KCCT Math		.696	
KCCT Reading			.550

Source: Program Review staff analysis of data provided by KDE.

To provide another measure of the degree to which CATS and ACT scores are related, students were ranked and then grouped into four equal groups—quartiles—based on their CATS scores. As Table 3.4 shows, about 55 percent of those who ranked in the bottom fourth based on CATS scores also ranked in the bottom fourth on the ACT. Less than 15 percent ranked in the top half of ACT scorers. More than 85 percent of those who scored in the top fourth on CATS scored in the top fourth on the ACT as well. About 75 percent of those in the second (from lowest) CATS quartile scored in the second or third ACT quartile. More than 80 percent of those in the third (from lowest) CATS quartile were in the third or fourth ACT quartiles.

Table 3.4
Student-level CATS Scores Grouped by ACT Score
(percentage of CATS-quartile students by ACT quartile)

	Lowest ACT Quartile	ACT Quartile 2	ACT Quartile 3	Highest ACT Quartile
Lowest CATS Quartile	55.6%	30.9%	9.6%	3.9%
CATS Quartile 2	12.0	40.4	34.0	13.7
CATS Quartile 3	2.1	14.3	35.0	48.6
Highest CATS Quartile	.7	2.1	11.4	85.8

Source: Compiled by Program Review staff from data supplied by KDE.

It is impossible to say for certain why there is a relatively strong correlation between students' ACT and CATS scores, but statewide ACT scores have not increased over the past few years and CATS scores have. The number of students taking the ACT over the past few years has increased. According to data from the ACT, approximately 9 percent more Kentucky students took the ACT in 2002 than in 1994. If the lack of increase in scores was simply due to an increase in overall participation, there should be similar patterns in other states that had increases in the percentage of students taking the exam. The data indicate, however, that of the

19 states that had at least a 5 percent increase in students taking the ACT from 1994 to 2002, 10 states showed an increase in state average ACT scores and 9 showed a decrease.

Another possible explanation is that more students who score poorly on the CATS assessment are now taking the ACT. Scores from mathematics, science, and social studies components of the CATS assessment were averaged for students who took the ACT between 1999-2002.⁴ The data indicate that there has been an increase in students who, on average, scored novice on the mathematics, science, and social studies portions of the CATS assessment taking the ACT, possibly lowering the average ACT score. In 1999, approximately 17 percent of the public school students who took the ACT scored novice; in 2002, 27 percent scored novice.

KCCT and CATS Are Valid and Reliable According to NTAPAA

KCCT and CATS are valid and reliable according to the National Technical Advisory Panel on Assessment and Accountability.

The National Technical Advisory Panel on Assessment and Accountability (NTAPAA) is a panel of nationally recognized testing experts who are paid through a contract with the Legislative Research Commission for advising the LRC and education officials. Members of the Program Review and Investigations Committee directed staff to determine whether CATS had been judged valid by NTAPAA. Apparently, the panel had not issued a document specifically stating that CATS was valid, but did so for this report.

NTAPAA was created in April 1998 in accordance with KRS 158.6454. The LRC is responsible for appointing NTAPAA members, and as of August 2003, the panel consisted of six experts in testing and measurement from various higher education institutions around the country. NTAPAA is charged with advising the Legislative Research Commission, the Kentucky Board of Education, and the Kentucky Department of Education on the following components of CATS:

- Development and implementation of CATS, including assessment and accountability components;
- Assessment design and achievement of goals;

⁴ The mathematics, science, and social studies components of the CATS assessment receive the same weight on the academic index. The reading component also receives the same weight. However, there was an insufficient number of ACT takers with CATS reading scores, so these scores were not included. The other CATS components receive less weight on the academic index and were not included.

- Validation studies;
- School report cards;
- School accountability index;
- Consequences for schools not meeting goals;
- Scholastic audits; and
- The formula for school accountability and improvement goals.

The complete text of NTAPAA's statement on the validity and reliability is available as Appendix F of this report. According to the document's summary:

There is substantial evidence supporting the validity and reliability of the Kentucky Core Content Test and the validity of the Commonwealth Accountability Testing System in the context of current implementation and use in the State of Kentucky. The information generated through this system provides valid and reliable indicators of student performance across schools and districts. Statistically significant changes in index scores over time should be considered valid indications that average student performance levels have changed over time. The evidence of student performance produced through KCCT and CATS constitutes a valid basis for rewarding or identifying for improvement schools and school districts to the degree provided in current statutes and regulations.

Chapter 4

Writing Portfolios and On-demand Writing

CATS uses writing portfolios and on-demand writing to evaluate students' writing skills. This chapter begins by comparing the characteristics and scoring of these assessments. The remainder of the chapter is devoted to writing portfolios—their compilation, scoring, and the accuracy of that scoring. Because portfolios account for 80 percent of a school's writing index, inaccurate scoring of portfolios can affect a school's overall academic index.

This chapter also explains the formal audit procedures KDE uses to assess the accuracy of schools' portfolio scores. KDE selects a sample of schools each year for auditing, choosing half of those purposefully according to set criteria and choosing the other half randomly. Selected schools must submit all portfolios for rescoring and any inaccurate scores are corrected. Apart from the adjustment of incorrect scores, there are no consequences to schools that inaccurately score portfolios. The results of the audits conducted for the past three years are presented and results are compared among scoring methods. Three recommendations are made in this chapter in an attempt to improve the accuracy of schools' portfolio scoring.

Background

The 1990 General Assembly directed that a performance-based assessment of writing for Kentucky's public school students be included in the system of assessment. In 1993, KDE and a committee of Kentucky English/language arts teachers designed the writing portfolio to teach students to write for multiple purposes, in multiple forms, and for a variety of audiences. Educators believed portfolios would be a means of assessment that directly reflected what students were learning because the content of the portfolios would come from students' classroom work. Kentucky is the only state that holds schools accountable based on

students' performance in writing portfolios and is one of only two states that requires students to create writing portfolios.¹

These educators placed an emphasis on developing broad writing abilities among Kentucky public school students with six criteria used to assess the quality of the students' writing:

- Awareness of the audience for and purpose of the writing;
- Development of ideas and support for them;
- Organization;
- Sentence structure and variety;
- Language (word choice and usage); and
- Correctness (spelling, punctuation, and capitalization).

Writing portfolio and on-demand writing are assessed in grades 4, 7, and 12.

The writing portfolio, which is assessed in grades 4, 7, and 12, is a compilation of the students' best writings, produced over an extended amount of time, with at least one piece from each of the following broad categories of writing:

- Reflective writing—a letter to the reviewer of the portfolio in which a student discusses his or her growth as a writer by reflecting on the pieces in the portfolio;
- Personal expressive writing—a narrative that focuses on one event in the life of the writer; a memoir that focuses on a relationship of the writer with a particular person, place, animal, or thing; or an essay that focuses on a central idea supported by a variety of incidents in the writer's life;
- Literary writing—a short story, poem, play, or script; and
- Transactive writing—writing produced for a variety of audiences and purposes in real world forms such as a letter, article, editorial, or speech.

The on-demand writing component of the CATS assessment is administered during the two-week testing window each spring to students in grades 4, 7, and 12. The on-demand assessment offers students a choice of two writing prompts. Each student selects one of the prompts and responds to it within a specific amount of time.

¹ Vermont has incorporated the use of portfolios in its public elementary and secondary school teaching regimen, but portfolio scores in Vermont's public schools do not play a part in that state's school accountability system. Portfolio assessment is used only as a guide to help schools improve instruction.

Writing portfolios and on-demand writing are given one of four scores: novice, apprentice, proficient, or distinguished. More than 60 percent of surveyed teachers agreed that these categories are appropriate.

On-demand writing and writing portfolios are each scored on a four-point scale. Scores are converted to a 140-point scale for incorporation into a school's index.

Novice	13 points
Apprentice	60 points
Proficient	100 points
Distinguished	140 points

In a survey administered by Program Review and Investigations Committee staff, more than 60 percent of teachers generally and 68 percent of teachers who have scored portfolios indicated that the performance categories novice, apprentice, proficient, and distinguished are appropriate categories to assign to their students' writing.

The writing index, produced from students' portfolio and on-demand scores, makes up 14.25 percent of the school's accountability index.

The scores assigned to the portfolios and on-demand writing assessment are then used to calculate the school's writing index. The writing portfolio score accounts for 80 percent of the writing index and the on-demand writing score accounts for 20 percent. The writing index accounts for 14.25 percent of a school's accountability index score.

Prior to 1998-1999, portfolios for all grades contained six entries of student writing. Since the 1998-1999 school year, grade-4 portfolios have contained four pieces, and grades 7 and 12 each have contained five pieces. The change was in response to KRS 158.6543(2), which required the Board of Education to reduce the amount of work required in the writing portfolio assessment.

Statewide, in both 2001 and 2002 approximately 800 students were exempted from the writing portfolio requirement and 340 students were exempted from the on-demand writing assessment.

According to KDE, all students in grades 4, 7, and 12 are required to complete the on-demand writing assessment and writing portfolio, except:

- Students with limited English proficiency who have been enrolled in a public school for fewer than two years;
- Foreign exchange students;
- Students expelled for behavioral reasons during the current or prior school year who are not receiving educational services;
- Students unable to participate due to medical reasons who have a signed doctor's statement; and
- Students not enrolled in a Kentucky public school for at least 100 instructional days prior to the testing date.

Kentucky Educators' Opinions of the Writing Portfolio

More than 70 percent of teachers indicated the amount of time needed to prepare portfolios is not appropriate for the benefit students receive.

In a survey administered by Program Review and Investigations Committee staff, Kentucky public school teachers were asked a series of questions regarding writing portfolios. Approximately 71 percent of teachers responded that the amount of time it takes for students and teachers to prepare portfolios is not appropriate to the benefit the students receive from doing them. The survey asked teachers, principals, and superintendents about the appropriateness of the weight assigned to each of the components of the accountability index. Forty-eight percent of teachers, 31 percent of principals, and 38 percent of superintendents indicated the writing portfolio received too much weight in the accountability index. When asked what changes they would make to the CATS assessment, more than 40 percent of teachers, 17 percent of principals, and 5 percent of superintendents responding to the question suggested that they would reduce the writing requirements or eliminate the writing portfolio.

A Comparison of Writing Portfolios and On-demand Writing

Correlations between writing portfolio and on-demand writing scores are consistent across grade levels, ranging from 0.40 to 0.45.

Both assessments are designed to measure a student's writing ability, although the tests do so in different ways. The on-demand writing assessment measures a student's ability to produce a written piece on a specific topic within a certain amount of time. The writing portfolio allows students to edit and revise their writing pieces and contains different types of writing produced over time. Student-level assessment scores indicate that writing portfolio scores are typically higher than on-demand writing scores. Comparisons were made of students' writing portfolios and on-demand writing scores over the past four years. Correlations between writing portfolios and on-demand writing scores are consistent across grade levels, ranging from 0.40 to 0.45. The correlations between students' writing portfolios and on-demand writing scores were 0.40 in grade 4, 0.45 in grade 7, and 0.43 in grade 12.²

It is unclear what the correlation between the types of writing assessments should be. Correlations of 0.40 to 0.45 indicate there is a positive relationship between the two measures of writing ability. The relationship is not a particularly strong one, however.

² A correlation coefficient, which can range from -1 to +1, measures the extent to which the values of two or more measurements go up and down together. A definition is provided on page 49 of this report.

It is possible that the two assessments are measuring different types of writing ability. Differences in scores could stem from the fact that students do not receive assistance in the on-demand writing assessment but can receive feedback from educators and other students on their writing portfolios. Another possibility is how the two types of writing are scored. School personnel score writing portfolios at the students' schools. CTB McGraw-Hill, the scoring contractor, scores the on-demand assessment.

Scores for writing portfolios were higher than for on-demand writing at all grade levels. The difference was less, however, for 12th graders.

Portfolios and on-demand scores can be broken down by students' grade levels and the scores assigned. Table 4.1 shows the results. Apprentice was typically the most common score, but portfolios had fewer students scoring novice and more scoring proficient. When examined by grade, a much higher percentage of 4th and 7th graders scored novice on the on-demand writing assessment than 12th graders. A higher percentage of 7th graders scored novice on the writing portfolio compared to 4th and 12th graders. Finally, few students scored distinguished on writing portfolios or on the on-demand writing assessment.

Table 4.1
Writing Portfolio and On-demand Writing Scores, 1999 to 2002
(percentage of students by score)

Grade		Novice	Apprentice	Proficient	Distinguished
4	On-demand	47.2%	46.5%	4.4%	.02%
	Portfolio	20.7	53.3	24.3	1.5
7	On-demand	41.7	49.3	6.7	.05
	Portfolio	47.3	39.0	11.3	1.1
12	On-demand	27.1	56.5	12.8	.06
	Portfolio	23.4	50.6	22.0	2.5

Source: Compiled by Program Review staff from data supplied by KDE.

Training for Scoring of Writing Portfolios

Teachers score writing portfolios at the school level so they can use what they learn from the assessments to improve classroom instruction. KDE provides training in an effort to ensure the accuracy of scoring and to help teachers align their instruction with the criteria for writing portfolios. All scoring sessions are centered on materials KDE develops, which are provided to scorers in the Kentucky Writing Portfolio Scoring Handbook. These materials include:

- Directions for conducting a scoring session;
- Forms needed for the scoring session;

- Benchmark portfolios, or portfolios that clearly reflect the desired components indicated by the guidelines for a particular performance level;
- Exemplar portfolios, which are used to support the benchmark portfolios by providing more samples of student performance at each performance level;
- High-end portfolios, or portfolios that demonstrate the upper end of each performance level, are used to determine where one performance level ends and another begins; and
- Quality control portfolios, which are portfolios that have already been scored by a KDE scoring accuracy team.

Scorer training sessions also include viewing a video produced by KDE that addresses frequent questions and concerns. Each school also receives a scoring training CD-ROM, and KDE provides additional portfolios that are used as training or validity portfolios at the local scoring session.

KDE personnel train regional writing consultants, who train writing cluster leaders, who train school-level portfolio scorers. Teachers reported receiving an average of four and a half hours of portfolio training annually.

KDE uses a trainer-of-trainers model to provide portfolio-scoring training. KDE personnel train regional writing consultants who, in turn, train writing cluster leaders at regional meetings. The cluster leaders train portfolio scorers at individual schools during training sessions that last three to six hours.

KDE requires that all members of a school's scoring team receive current scoring training for the year in which they are to score portfolios. A minimum of three hours of annual training is required of all school personnel who score portfolios, but six hours is recommended. According to the Program Review staff survey of Kentucky teachers, teachers who score portfolios receive an average of around four and a half hours of training annually.

The majority of teachers found the high-end portfolios helpful. About half did not find the training CD-ROM or video helpful.

The survey asked teachers their opinions of the writing portfolio scorer training. Approximately 90 percent indicated the training delivered by their school's cluster leader was helpful and 85 percent indicated that they utilized the high-end portfolios during their training sessions. When asked to explain which aspects of the portfolio scoring training were particularly helpful, almost half the teachers indicated that the quality control, benchmark, and high-end portfolios were helpful, as were scoring sample portfolios and group discussions.

When asked to explain which portions of the writing portfolio scoring training they would change, approximately 10 percent of the teachers who answered the question indicated that new benchmarks portfolios were needed because the same samples had

been used for too many years. This was the most common suggestion provided. When asked if they used the scoring training CD-ROM provided by KDE, almost half indicated that they did not. Fifty-one percent of teachers indicated that they found the video portion of the portfolio scoring training not very helpful or not helpful at all.

Almost half of the teachers responding to the survey said that “a few” or “many” teachers had biases that affected scoring of portfolios.

Teachers and principals were asked if the teachers who scored portfolios had biases that affect scores. A majority of principals, 55 percent, responded that teachers did not have biases that affected scores. Twenty percent of principals indicated that a few teachers had biases that affected scoring. Only 6 percent responded that many teachers had such biases. Almost half the teachers who responded said that teachers had biases that affected the scoring of portfolios, including 34 percent attributing scoring biases to a few teachers and 14 percent who said many teachers were biased.

Approximately 30 percent of teachers report scoring their own students’ portfolios.

According to the responses of both teachers and principals, slightly more than 30 percent of teachers are scoring their own student’s portfolios. When asked to explain how the scoring was affected, the responses volunteered most often by teachers were that it was hard to be objective when it came to their own students, that they had personal biases that affected their scoring, and that they based their scoring on the amount of work students had put into the portfolio pieces rather than the piece itself.

When asked if their schools had enough teachers scoring portfolios, 77 percent of teachers and 89 percent of principals indicated that their schools had enough teachers scoring writing portfolios. When asked to explain what steps could be taken to improve writing portfolio training, the responses volunteered most often were that training should be required of all teachers in the school, that the benchmark and exemplar pieces provided in the training material should be changed, and that there should be more opportunity to practice portfolio scoring throughout the year.

Recommendation 4.1

KDE should work with schools and districts to reduce, as much as possible, the practice of teachers scoring their own students’ portfolios. KDE should survey teachers to determine how their portfolio scoring training can be improved. KDE should regularly replace benchmark portfolios with new samples. KDE should encourage schools to provide teachers with more opportunities to practice scoring writing portfolios.

Scoring Options

KDE provides the materials and guidelines for grading writing portfolios, and schools decide independently which procedural option they will use to score their portfolios. There are six procedural options that a school can choose for scoring portfolios:

Option 1 - Double Blind Scoring. Each teacher scores a certain number of portfolios in a group setting, noting the score of each portfolio with a self-stick note placed on the back of the portfolio. Each scored portfolio becomes part of a central group and is scored by a second teacher who does not know the initial score. The second score is noted on the front of the self-stick note. The two scores are compared. Portfolios with different scores are pulled, discussed, and a final score is assigned through consensus.

Option 2 - Individual Scoring/Blind Second Scoring. Individual teachers working independently complete the initial scoring. Scorers then meet as a group and a second teacher who does not know the initial score scores each portfolio. The two scores are compared. Portfolios with different scores are pulled, discussed, and assigned scores through consensus.

Option 3 - Individual Scoring/Selected Blind Second Scoring/Reliability Check Session. Portfolios are scored independently. Scorers select portfolios that presented difficulties in scoring and represent diverse performance levels. These portfolios are brought to a blind second scoring session for review, discussion, and consensus scoring.

Option 4 - Individual Scoring/Selected Blind Second Scoring. Portfolios are scored independently. Scorers select portfolios that present difficulties in scoring. In a blind second scoring session, these portfolios are reviewed and discussed. Final scores are assigned through consensus.

Option 5 - Individual Scoring/Selected Second Scoring/Informal Pair Discussion. Portfolios are scored independently. Portfolios that were difficult to score are given to a second scorer to be rescored independently. The two scorers meet to discuss any differences and assign a final score together.

Option 6 - Individual Scoring with Informal Support/Discussion as Necessary. Individual scorers working independently score portfolios. As necessary, scorers meet to informally discuss portfolios that were difficult to score and assign a final score together.

The majority of schools use the double blind method for scoring writing portfolios.

KDE does not require the use of a particular option, but discourages option 6 because it does not provide any group support or feedback from fellow scorers. The scoring options of schools have changed somewhat over the past four years, but the most common method of scoring portfolios has been double blind scoring (option 1). According to KDE, approximately 85 percent of Kentucky schools that indicated their scoring option used the double blind method in 2001-2002. The percentage of schools using this scoring method has increased since 1998-1999.

The survey asked teachers who indicated they were portfolio scorers to indicate how often conflicting scores were initially assigned to portfolios. Of the approximate 800 teachers responding to the survey, about 300 indicated they scored writing portfolios and answered this question. The responses ranged from 1 percent of the time to 95 percent of the time, with different scores being assigned, on average, approximately 18 percent of the time. Staff grouped the responses into four categories, which are listed in Table 4.2. More than half of those responding said initial disagreement was rare—10 percent or less of the time. About a fourth, however, answered that initial scores were different more than 20 percent of the time. Approximately three-fourths of teachers indicated that the scoring option their school utilized yielded accurate writing portfolio scores.

Table 4.2
Percentage of Time Teachers Report
That Initial Writing Portfolio
Scores Do Not Agree*

Teacher's Response Was:	% of Respondents
From 0 to 10%	53.2%
From 11 to 20%	20.7%
From 21 to 30%	15.7%
More than 30%	10.4%

* The question was "About what percentage of the time do the first two scores given to a portfolio by scorers working independently NOT agree?"
Source: Compiled by Program Review staff from the survey of teachers.

Monitoring of Portfolio Development and Scoring

Portfolio development and scoring are monitored in several ways. KDE distributes an administrative code to all schools describing the limits on a teacher's comments and modifications of students' portfolio entries. When the scores assigned by the school are submitted to KDE, the principal must submit a signed assurance statement confirming that appropriate portfolio development procedures were followed. Each student's portfolio includes a signed statement attesting that the student completed all entries. Finally, KDE performs a writing portfolio audit that monitors portfolio scoring.

Selection of Schools for Audit of Writing Portfolios

KDE uses a formal audit procedure of portfolio scoring.

KDE utilizes a formal audit procedure to increase the probability that portfolios are scored accurately. The portfolio audit is a review of local scoring accuracy and is used to monitor statewide scoring patterns and to adjust scores for schools that score portfolios inaccurately. KDE uses the audit to monitor the accuracy of scoring to help plan portfolio training and to allocate resources. KDE also uses the audit to correct inaccurate scores and to verify exceptional gains schools made in their CATS scores. Any school that is selected to participate in the annual audit must submit all portfolios for rescoring.

Approximately half the schools in the portfolio audit are chosen because of above-average changes in scores. The remaining schools are chosen at random.

Approximately 100 schools are selected for auditing each year. Since 1996, KDE has identified schools to be audited using two methods. Some schools are selected because they rate highly on an index that indicates unusual performance on writing portfolios. This is defined as the purposeful sample selection. The remaining schools to be audited are chosen randomly. KDE regulations specify that schools cannot be included in the purposeful sample two years in a row. A school may, however, be in the purposeful sample one year and be included in the random sample the following year.

For the purposeful selection, an index is created by subtracting the writing portfolio score from two years ago from the portfolio score of last year.³ This number is added to a calculation of last year's academic index less last year's portfolio score. The following formula is used for the purposeful selection:

(Writing Portfolio Score for Last Year - Writing Portfolio Score for Two Years Ago) + (Academic Index for Last Year - Writing Portfolio Score for Last Year)

In other words, a school ranks highly based on this measure if its writing portfolio scores changed significantly, up or down, and its overall academic index did not. Schools are then ranked based on the formula. Schools with the highest scores on the selection index are selected for inclusion in the audit. This formula identifies schools with unusual scores, but does not ensure a representative sample of grades and schools. More elementary schools and fewer high schools are selected because there are many more elementary schools. Smaller schools are also more likely to be selected because these schools have fewer students, so their scores are more likely to fluctuate from year to year. For example, in 2002 only 81 students were included in the 12th-grade audit. After the purposeful audit schools have been selected, a random sample is chosen from the remaining schools.

Beginning in 2002-2003, schools with the largest differences in actual and predicted portfolio scores will be included in the purposeful audit.

Beginning with the audit of the 2002-2003 scores, KDE will change its audit selection procedure and begin using simple linear regression—a statistical technique—to select a purposeful sample that is more representative of grades and schools. Each school's academic index, minus the writing portfolio component, will be used to estimate what the school's writing portfolio performance should be. Schools that have relatively high values on the academic index, excluding the writing portfolio component, should have relatively high writing portfolio scores too. Similarly, schools with relatively low academic index scores, excluding the writing portfolio component, should have relatively low writing portfolio performance. Schools will then be ranked based on the differences between their actual writing portfolio scores and their estimated scores. Schools with the largest differences will be selected for inclusion in the purposeful sample. Two-thirds of the schools selected will have writing portfolio scores higher than estimated and one-third will have scores lower than estimated. After the purposeful selection process is completed, the remaining schools

³ The current year's scores cannot be available until after the audit process is completed.

included in the audit will be selected randomly. Approximately 100 schools will still be audited each year.

Auditing of Writing Portfolios

Writing scorers from the scoring contractor, CTB McGraw-Hill, perform the audits of writing portfolios.

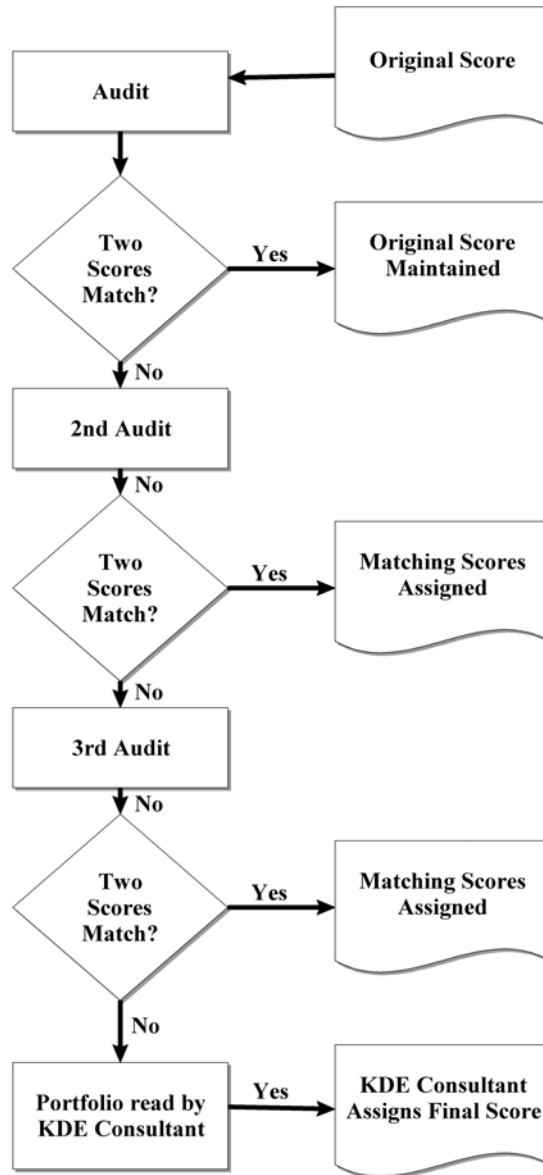
Writing scorers from the scoring contractor, CTB McGraw-Hill (CTB), perform the audits of writing portfolios.⁴ Writing portfolio consultants from KDE and CTB train all audit scorers using the same procedures and materials used to train scorers in Kentucky during the school year. Figure 4.1 shows how the audit procedure works and how, if necessary, new scores are assigned to portfolios. Portfolios are packaged within grade levels and an audit scorer provides a score for each portfolio. Audit scorers are not aware of any previously assigned score(s). For each portfolio, the auditor's score is compared to the original score assigned by the school. If the scores agree, the original score is maintained. If the scores do not agree, a second auditor scores the portfolio. After the second audit score is assigned, the three scores are compared. If two of the three scores agree, the score agreed on is assigned as the final score. If the three scores differ, the portfolio is scored by a KDE or CTB consultant. Any two scores that agree of the four scores are assigned as the final score. If there are still no two scores that agree, a final KDE consultant scores the portfolio. Previous scores are reviewed and a final score is assigned.

In the 2002 audit, approximately 36 percent of the first audit scores did not agree with the school's original score and had to be rescored by a second auditor. Only 2 percent of the total number of portfolios audited had to be scored by a third auditor and no portfolios had to be scored by a KDE consultant.

After selected portfolios have been audited, the scores assigned during the audit replace the original scores and the school's writing portfolio index is recalculated. The school's accountability index is then recalculated with the new writing portfolio index. The writing portfolio index makes up 11.4 percent of a school's overall accountability index.

⁴ Prior to the 2002-2003 writing portfolio audit, Data Recognition Corporation performed Kentucky's portfolio audits. The audit procedure was similar to the one explained above for CTB McGraw-Hill.

Figure 4.A
Writing Portfolio Audit Process



The Validity of the CTB Audit

KDE and CTB take several steps to better ensure the accuracy and reliability of audit results.

KDE and CTB take several steps to better ensure that audit results are reliable and provide the most accurate and consistent information. First, all scorers who participate in the auditing process must demonstrate a certain level of scoring accuracy on sets of portfolios for which scores have been predetermined by the Kentucky Writing Advisory Committee. Second, CTB lead scorers read and rescore 20 percent of all audited portfolios. If scoring discrepancies arise, discussion and resolution occur immediately. Scores assigned by both the lead scorer and the regular audit scorer are documented to compare to the original score to determine the level of internal agreement between scorers. KDE consultants and CTB scoring directors also conduct a similar consistency check by reading and rescoring approximately 20 percent of the portfolios scored by the lead scorers. Third, the Scoring Accuracy Assurance Team, a group of Kentucky scorers who have demonstrated a high level of scoring accuracy on a KDE-administered scoring verification process, conduct a separate audit review. They score a random sample of 20 percent of audited portfolios to verify the quality of CTB scorers. Finally, to provide continual retraining, two quality control portfolios—portfolios with scores predetermined by the Kentucky Writing Advisory Committee—are read and scored by each auditor daily.

Audit Results

The portfolio score assigned during the audit is the final score assigned to the portfolio, and any changes in scores are adjusted in the school's writing portfolio index.

As shown in Table 4.3, agreement among the school's scores and scores assigned during the audit have been highest for fourth-grade portfolios. Grade-12 portfolios showed the lowest levels of agreement over the past three years. The agreement rate is the percentage of students for which the original score and the auditors' score is the same. Depending on year, grade level, and whether the school was chosen for a purposeful or random audit, agreement rates range from about 60 percent to more than 80 percent. From another perspective, this means that the original and audited portfolio scores do not agree from less than 20 percent to almost 40 percent of the time. There does not appear to be a consistent trend toward higher agreement rates over the past three years. As would be expected, scores from schools that were randomly selected for audit typically showed higher levels of agreement than scores from purposefully selected schools.

Over the past three years, scores that were changed, for both randomly and purposefully selected schools, were lowered usually.

Most often, the percentage of students that scored novice was increased significantly and the percentages that scored proficient and distinguished were reduced significantly. For instance, the 1999-2000 audit resulted in about 50 percent more novice portfolios than originally reported by the audited schools. The 2001-2002 audit resulted in approximately 36 percent more novice portfolios. There were about 40 percent fewer proficient portfolios in 1999-2000 than had been reported by the audited schools and 35 percent fewer proficient portfolios in 2001-2002.

Table 4.3
Agreement Rates Between Original Scores and Scores
Assigned During Audits for Writing Portfolios
(number of students in parentheses)

Year	Grade	Purposeful	Random
1999-2000	4	76.9% (1,963)	80.2% (1,958)
	7	71.5% (1,637)	76.4% (2,336)
	12	67.8% (425)	60.1% (1,565)
2000-2001	4	78.8% (1,960)	82.6% (2,076)
	7	68.1% (238)	83.0% (2,490)
	12	80.0% (921)	75.5% (2,072)
2001-2002	4	80.5% (1,279)	82.6% (1,798)
	7	75.9% (3,236)	73.1% (1,940)
	12	60.5% (81)	73.5% (1,468)

Source: Compiled by Program Review staff from data supplied by KDE.

Table 4.4 presents, by scoring option, the percentage of original scores that were changed by auditors over the past four school years. Because option 1 (double blind scoring) is used most often, its accuracy changes the least from year to year. For most options for most years, the share of scores changed is about 20 to 25 percent. Option 6 (individual scoring with informal support/discussion as necessary) seems to be the least accurate method. Due to data limitations, it was not possible to determine if there were similarities among schools that chose a specific scoring method, such as small schools choosing a particular method, for example. Based on three years of audits, it is not obvious that any option stands out for accuracy, except for option 6 appearing worse than others.

Table 4.4
Percent of Scores Changed in Writing Portfolio Audit
by Scoring Method, 1998-1999 to 2001-2002

Scoring Option	Year:	1998- 1999	1999- 2000	2000- 2001	2001- 2002
	Percent Changed				
1. Double Blind Scoring		24.1	22.5	19.2	23.2
2. Individual Scoring/Blind Second Scoring		22.6	21.6	20.0	26.1
3. Individual Scoring/Selected Blind Second Scoring/Reliability Check Session		25.4	33.2	21.1	23.3
4. Individual Scoring/Selected Blind Second Scoring		24.9	None audited	None audited	None audited
5. Individual Scoring/Selected Second Scoring/Informal Pair Discussion		15.3	28.9	24.1	17.5
6. Individual Scoring With Informal Support/Discussion as Necessary		31.8	27.5	21.4	42.2

Source: Compiled by Program Review staff from data supplied by KDE.

Schools chosen for an audit, whether randomly or purposefully, face no consequences for high numbers of improperly scored portfolios other than the replacement of the original scores with the audited scores in their accountability index. A school chosen for a purposeful audit also cannot be chosen intentionally again the following year. If a purposefully audited school had a high number of inaccurate scores one year, there would be no follow-up the next year to see if scoring accuracy had improved unless the school was randomly selected, which would be unlikely. Given that accuracy of portfolio scoring does not appear to be improving significantly,

it may be necessary to increase the accountability for portfolio scoring. Portfolio scorers are chosen at the school level and KDE has no way of tracking which scorers are the least likely to have their scores changed during auditing. KDE should consider implementing such a system to ensure that portfolios are scored by the most accurate and most experienced scorers possible. KDE should also consider establishing consequences—rewards, sanctions, or both—for schools to encourage better scoring.

Recommendation 4.2

KDE should consider implementing a system to track the performance of portfolio scorers.

Recommendation 4.3

KDE should consider establishing consequences for schools that have low portfolio audit agreement rates, such as having audit agreement rates reflected in the school's accountability index. KDE should consider reauditing schools that had a high number of scoring inaccuracies the prior year to ensure that scoring accuracy has improved. KDE should also consider increasing the number of schools randomly selected for audits so that the risk of facing consequences would encourage schools to score more carefully.

Chapter 5

Kentucky's Compliance With the No Child Left Behind Act

The No Child Left Behind Act (NCLB) increases demands on states for educational standards and assessment.

Under the federal No Child Left Behind Act,¹ states are required to define a proficiency level that all students must reach by the 2013-2014 school year. States must test students in reading and math annually, and science in specified grades. Each state must implement a statewide accountability system that establishes sanctions and rewards for schools based on students' performance against the state's standard. Other provisions require that states analyze and report test results in new ways, provide supplemental services for students who attend low-performing schools, and require teachers to be "highly qualified."

States not complying with No Child Left Behind (NCLB) risk the loss of federal funds for elementary and secondary education. In 2002, Kentucky state government and local school districts spent about \$3 billion for elementary and secondary education, of which about \$476 million was federal funding.

Kentucky's Commonwealth Accountability Testing System (CATS) predates NCLB. CATS is similar to the accountability system required by NCLB, but does not satisfy all of NCLB's requirements. This chapter begins with a brief description of the status of Kentucky's compliance with NCLB and then discusses some of the key requirements of NCLB dealing with the accountability system, challenging standards, academic assessment, graduation rates, failing schools, safe schools, and qualifications of teachers and paraprofessionals. As each NCLB requirement is discussed, it is compared to the requirements of CATS. Information is also provided on other states' progress toward meeting NCLB's requirements. Kentucky educators' opinions on implementing two NCLB requirements are presented. The chapter ends with a discussion of the costs of implementing NCLB.

¹ No Child Left Behind Act of 2001, Public Law No. 107-110, 115 Stat. 1425

Remaining NCLB Compliance Issues

The United States Department of Education approved the basic elements of Kentucky's NCLB accountability plan on June 18, 2003, contingent upon the resolution of remaining issues.

No Child Left Behind requires any state applying for federal education grants to submit a plan that satisfies all of NCLB's requirements. Kentucky submitted its NCLB accountability plan to the United States Department of Education for approval. The department approved the basic elements of the plan on June 18, 2003, noting that the plan satisfied most, but not all, of NCLB's requirements. Like other states, Kentucky still must address some remaining issues to comply with NCLB.

The department accepted Kentucky's plan contingent upon the resolution of inconsistencies in the areas described below:

- Calculating adequate yearly progress in math and reading;
- Adding math and reading testing in additional grade levels;
- Assessing only those students enrolled in the school for a full academic year;
- Calculating graduation rates;
- Identifying schools for improvement and issuing school reports in a timely manner; and
- Incorporating accountability for districts in addition to schools.

The Kentucky Board of Education has decided to implement both the NCLB and CATS accountability systems.

At the June 2003 Education Assessment and Accountability Review Subcommittee meeting, Kentucky Department of Education (KDE) staff and the chair of the Kentucky Board of Education discussed possible solutions to the remaining issues to be addressed under Kentucky's No Child Left Behind plan. At its meeting on August 7, 2003, the Board of Education decided to retain CATS unchanged and incorporate the additional requirements of NCLB. The two accountability systems will operate concurrently.

Kentucky and other states are working out remaining issues with the U.S. Department of Education regarding implementation of NCLB.

Like Kentucky, other states are still working to implement No Child Left Behind and address issues raised by the U.S. Department of Education. Although all 50 states' NCLB accountability plans were conditionally approved by June 2003, states' progress toward implementation varies. Appendix G summarizes states' status with several provisions of NCLB.

Accountability System

Some differences between NCLB and CATS make them difficult to integrate.

Some of the major components of the No Child Left Behind accountability system are similar to those of CATS [§1111(b)(2)].²

² The relevant subsection numbers of the No Child Left Behind Act are included in brackets throughout the chapter.

Both systems include a long-term goal of proficiency to be achieved by 2014 and both use periodic reviews of schools' progress resulting in sanctions and rewards intended to hold school districts and schools accountable for student achievement. There are, however, differences that make it challenging to integrate the two systems. NCLB and CATS calculate their system starting point, or baseline, differently and use different measures of success. CATS assesses students in numerous subjects; NCLB focuses on students' performance in reading and math and requires improvement by specific subgroups of students. The two systems incorporate different timetables for evaluating schools and distributing results. Some of the key differences between the systems are discussed below.

NCLB requires all schools and districts to make "adequate yearly progress."

No Child Left Behind requires that each state implement a single, statewide accountability system that will ensure that all school districts and all public schools make "adequate yearly progress" as measured by students' performance on the yearly assessments. CATS requires schools to reach their index score goals by the end of each two-year cycle.

Baseline Calculation

Baselines are to be calculated differently for NCLB and CATS.

An important difference between the No Child Left Behind and CATS systems is the calculation of the baseline against which a school's progress is compared. Under CATS, each school's individual baseline is computed from its 1999-2000 test scores. NCLB requires all schools in the state to have the same baseline starting point. The NCLB baseline is based on data from the 2001-2002 school year and equals the higher of the following: 1) the percentage of proficient students in the state's lowest achieving group, or 2) the percentage of proficient students in the school ranked at the 20th percentile of all schools. Under NCLB, the baselines are to be calculated separately at the elementary, middle, and high school levels. The baselines are then used to determine adequate yearly progress for specific student subgroups. Under the two-dimensional system approved by the Kentucky Board of Education, it will be necessary to calculate new baselines for NCLB and compare schools' performance according to both baselines.

Proficiency of All Students

NCLB requires states to set statewide annual goals in math and reading that identify a percentage of students required to attain proficiency. With the exception of the “safe harbor” provision, a school has failed to make adequate yearly progress if a subgroup does not score at or above the annual goal.

After establishing the starting point under No Child Left Behind, each state is to set statewide goals, or annual measurable objectives, for mathematics and reading. The objectives must identify a single percentage of students who are required to meet or exceed the proficient level for each of four specific groups of students: 1) economically disadvantaged students; 2) students from major racial and ethnic groups; 3) students with disabilities; and 4) students with limited English proficiency. The objectives increase in equal increments over the period covered by the state’s timeline, and all students are to reach proficiency by 2014. A school has made adequate yearly progress if each of the four specified subgroups of students meets or exceeds the state’s objectives for that year.

NCLB does create a “safe harbor” exception if the percentage of students or a subgroup of students scoring proficient increases by at least 10 percent. Consistent with NCLB’s exception, at its August 7, 2003, meeting the Kentucky Board of Education approved a policy of allowing an exception for schools receiving federal Title I funding. If a Title I school does not meet the reading or math requirement for adequate yearly progress, the school will be considered to have made progress if 1) within the school or subpopulations that did not meet federal requirements, the percentage of students scoring below proficient is reduced by at least 10 percent; and 2) the students in the same subpopulation demonstrate improvement on the academic index (elementary and middle schools) or on the graduation rate (high schools).

Reducing Achievement Gaps. This is one of the primary goals of No Child Left Behind. NCLB requires schools to show student progress toward achievement, not just among the student population as a whole, but for racial and ethnic minority students, economically disadvantaged students, students with disabilities, and students with limited English proficiency.

Many Kentucky principals and superintendents expressed concern with their schools’ ability to close some achievement gaps.

In the surveys conducted for this study, Program Review staff asked principals and superintendents whether their schools could close the gender, disability, English language proficiency, income, and race/ethnicity gaps. Many superintendents and principals, around 40 percent of each group, reported that their districts and schools would not be able to close the disability achievement gaps. A majority of principals responded that their schools could not close the income achievement gap by 2014.

In response to an open-ended question that specifically addressed closing achievement gaps, many educators expressed concern about closing the gap between low-income students and their more affluent peers and between students with disabilities and those without disabilities. Many indicated it would be impossible to close the disability achievement gap because educators cannot change students' IQs or lessen a student's disability regardless of the testing accommodations they provide. Many also expressed concerns whether schools can address the conditions under which low-income students live outside school.

Only 16 States Hold Schools Accountable for Subgroup Performance

Most states are not holding districts and schools accountable for subgroup performance.

Most states do not hold districts and schools accountable for subgroup performance, nor do they indicate test scores for all the required subgroups. Sixteen states hold districts and schools accountable for subgroup performance, 18 states test 95 percent of students in a subgroup, and only 7 report results by the required subgroups.

Kentucky reports results for all the required subgroups but exclusion rates have been relatively high for students with limited English ability and students with disabilities. Kentucky must still comply with NCLB by testing 95 percent of students in subgroups and holding schools and districts accountable for subgroup performance based on adequate yearly progress.

Assessments in Additional Subjects

NCLB requires an academic indicator in addition to math and reading. Kentucky plans to use CATS assessments already administered in other subjects.

No Child Left Behind focuses on students' performance in reading and math. Kentucky's system holds schools accountable for students' performance in several additional subjects: science, social studies, arts and humanities, writing, and practical living and vocational studies. NCLB does require states to use another academic indicator in addition to the math and reading assessments to determine adequate yearly progress. States are to choose that indicator. Kentucky proposes to use the CATS assessments in science, social studies, arts and humanities, and practical living and vocational studies as the additional indicator. Math and reading scores would still count more under NCLB, but schools and districts would be required to show improvement on this additional indicator for the purpose of evaluating yearly progress. All the academic indicators must be measured separately for each of the four NCLB subgroups.

Measuring Success

NCLB and CATS measure success differently.

Another significant difference between No Child Left Behind and CATS is the way the two systems measure success. Schools' and districts' progress are reviewed every year under NCLB. With CATS, assessments are administered each year but the index is calculated every two years. Additionally, NCLB and CATS focus on slightly different aspects of a school's performance in determining success or failure. NCLB identifies the primary indicator of success as the percentage of students who have attained proficiency. Improvement in student assessment results that do not amount to an increase in the number of students who are proficient will not meet a school's NCLB goals. Under CATS, a school's accountability index reflects the school's scores in every category; therefore, a general increase in scores will be reflected in an increased accountability index, even though there may have been little or no increase in the number of students who are proficient.

Concurrent Accountability Systems May Cause Confusion

Using both systems concurrently could result in confusion by creating conflicting classifications for the same school.

The concurrent use of both accountability systems may result in conflicting classifications of the same school or district. A school or district could be classified as failing under NCLB but meeting its goal under CATS, which could result in confusion for educators, parents, students, and the public. For example, a school could fail to meet its annual goal under NCLB because one subgroup did not make adequate progress, yet it could have successfully met its goal under CATS due to an overall increase in scores.

The timing of the release of assessment results may also create confusion. Under No Child Left Behind, schools and districts must receive NCLB test results by the end of the summer so that if a school is failing, parents will have the opportunity to transfer their children to a better performing school before the beginning of the next school year. A few months later, in October, the same school could be classified as meeting its goal under CATS.

Challenging Standards

NCLB requires states to adopt challenging academic content and student achievement standards in reading, math, and science.

No Child Left Behind requires each state to adopt challenging content standards and student achievement standards now in, at a minimum, math and reading [§§1111(b)(1)]. No later than the 2005-2006 school year, states must adopt standards in science as

well. The standards must apply to all public elementary and secondary students in the state. A state with standards in place may revise them to comply with NCLB requirements rather than adopt new standards.

Kentucky's current reading and math standards comply partially with NCLB's requirements.

Kentucky's current reading and math content standards partially comply with NCLB's requirements. NCLB calls for separate standards for each grade level; Kentucky uses a single standard for each school level—elementary, middle and high school. Kentucky intends to continue using its existing standards. The U.S. Department of Education has approved this portion of Kentucky's plan and is allowing flexibility for states like Kentucky that have standards for multiple grades.

NCLB gives discretion to states to determine the definition of proficiency. In August 2003, the Kentucky Board of Education decided that "proficient" under CATS will correspond to "proficient" under NCLB.

NCLB requires the states' academic achievement standards to be aligned with their academic content standards and to describe three levels of achievement: basic, proficient, and advanced. It is left to the discretion of the states to define "proficient" for NCLB purposes.

Under CATS, eight achievement standards, ranging from nonperformance to distinguished, have been established in reading and math.³ In August 2003, the Kentucky Board of Education decided that "proficient" under CATS will correspond to "proficient" under NCLB.

Math and Reading Standards Need To Be Developed in Half the States

Math and reading standards need to be developed in almost half the states.

As indicated in Table 5.1, math and reading standards need to be developed in almost half the states. NCLB requires each state to adopt challenging content standards and student achievement standards in, at a minimum, math and reading for grades three through eight. No later than the 2005-2006 school year, states must adopt standards in science for elementary, middle, and high school levels.

³ The four main performance levels are novice, apprentice, proficient, and distinguished. Novice and apprentice are each further divided into the levels nonperformance, medium, and high in reading, mathematics, science, and social studies for a total of eight achievement standards.

Table 5.1
Where States Stand: NCLB Standards*

NCLB Standard	States Adopting Standards	Kentucky (Yes, Partially, No)
Reading Standards	27	Partially
Math Standards	27	Partially
Science Standards	48	Yes

*Includes District of Columbia.

Source: Compiled by Program Review Staff using an Education Commission of the States database.

States Lower the “Proficiency” Bar

States’ standards and assessments vary, and some have very different definitions of proficient. Louisiana, Colorado, and Connecticut have set lower proficiency levels to reduce the number of schools that would be labeled as failing under NCLB.

States’ standards and assessments vary, and some have very different definitions of proficient. To reduce the number of schools that would be considered to be failing under NCLB, some states have redefined “proficiency” and lowered their expectations of student performance. For example:

- Louisiana will consider students proficient under the federal law if they score at the state’s “basic” achievement level. To avoid confusion, Louisiana renamed the “basic” category “mastery.”
- Colorado students who score at the “partially proficient” level on the state test will be considered proficient under NCLB.
- Connecticut has set a new proficient level to satisfy NCLB and decrease the number of schools that would be labeled as failing under the federal legislation. This means that Connecticut students will be deemed proficient under the federal system even if they do not meet the state’s performance goals in reading and math.

Academic Assessment

NCLB requires states to implement “high-quality, yearly, student academic assessments” in math, reading, and science.

No Child Left Behind requires states to implement “high-quality, yearly, student academic assessments” in math, reading, and science [§1111(b)(3)]. NCLB establishes criteria the assessments must meet and dictates the grade levels and subjects in which they are to be administered. These assessments are to be used as the primary means of determining the yearly performance of the states’ school districts and schools and whether they have made adequate yearly progress.

NCLB requires that, initially, reading and math assessments must be administered every year, not less than once during grades 3 to 5,

once during grades 6 to 9, and once during grades 10 to 12. Beginning with the 2005–2006 school year, math and reading assessments must be administered every year in each of grades 3 to 8 and once in grades 10 to 12. Science assessments are not required until the 2007–2008 school year. Thereafter, they must be administered yearly, at least once during grades 3 to 5, once during grades 6 to 9, and once during grades 10 to 12.

Many of the tests required by NCLB are already administered under CATS, but Kentucky will have to add testing in math in grades four and seven, and in reading in grades five and eight.

Many of the tests required by NCLB are already administered under CATS. CATS already requires testing in science in grades 4, 7, and 11, which satisfies the NCLB science requirements. However, Kentucky does not test students in math in grades four and seven, or in reading in grades five and eight.

To comply with NCLB, both the Kentucky Board of Education and the National Technical Advisory Panel on Assessment and Accountability have recommended using a norm-referenced test that can be expanded as necessary to ensure coverage of both Kentucky’s core content and the NCLB performance standards. The Kentucky Board of Education proposes to administer the CTBS-5 reading and math tests, supplemented with open-response and multiple-choice items as needed, to cover the core content as well as satisfy the NCLB testing requirements.

In addition to the yearly assessments, NCLB requires states to participate in the National Assessment of Educational Progress (NAEP) biennial academic assessments of reading and math in grades four and eight [§1111(c)(2)]. Kentucky has participated in NAEP assessments and intends to continue to do so.

Thirty States Need To Add Assessments in Reading and Math

Most states are in the process of creating and implementing new assessments to fulfill NCLB requirements.

Most other states are in the process of creating and implementing new assessments to fulfill NCLB requirements. Twenty states have already produced the necessary reading and math assessments required in grades three through eight under NCLB. Approximately two-thirds of the states have met the NCLB one-time science assessment requirement in grades 3 to 5, 6 to 9, and 10 to 12.

Inclusion of Students in the Assessments

NCLB requires all students to be assessed, including disabled students and students with limited English proficiency.

Pursuant to No Child Left Behind, all students are required to participate in the assessments, including students with disabilities and students with limited English proficiency. Reasonable

adaptations and accommodations should be provided for both groups of students. Students with limited English proficiency may take assessments in another language or form if that is the most likely method to yield accurate results. However, reading assessments written in English must be used with any student who has attended school in the United States for three or more consecutive school years. Under limited circumstances, a local school district may continue to administer assessments to the student in a foreign language for two additional years.

Kentucky's regulations exempt students with limited English proficiency from CATS assessments.

Kentucky includes disabled students in CATS assessments, but students with limited English proficiency who have attended an English-speaking school for less than two full years are currently exempted from the assessments. KDE intends to amend current regulations to comply with NCLB.

NCLB requires annual assessments of English proficiency. Kentucky will implement annual testing.

NCLB requires that school districts must also administer an annual assessment of English proficiency to all students with limited English proficiency [§1111(b)(7)]. The assessment must measure students' oral language, reading, and writing skills in English. Kentucky has used English proficiency examinations to determine whether testing accommodations were necessary, but had no policy in place to administer such exams yearly. KDE proposes to modify the relevant regulation to require annual assessments.

NCLB requires that transient students be excluded from accountability decisions unless they have attended the same school or district for one full academic year. CATS includes transient students in assessment and accountability decisions.

NCLB requires that transient students be included in the assessments only if they have attended schools within a single school district for one full academic year. The performance of students who have attended more than one school in the school district in any academic year is to be used only in determining the progress of the local district, not the school.

Under CATS, Kentucky includes all transient students enrolled in the school on or before the first day of the testing window regardless of the length of time the student has been enrolled. To comply with NCLB, Kentucky proposes to continue to assess any student who is enrolled on the first day of the testing window, but exclude the scores of those students who have not been enrolled in the same school or district for a full academic year from the NCLB accountability scores. NCLB leaves it up to the states to define a full academic year as long as the definition does not exceed 365 days. KDE proposes to define enrollment for an academic year as enrollment for 100 instructional days and enrollment on the first day of the testing window.

Graduation Rates

NCLB requires the inclusion of graduation rates in determining adequate yearly progress. Kentucky will exclude the number of students with certificates of completion from the number of graduates.

No Child Left Behind mandates the inclusion of graduation rates as a component of adequate yearly progress for high schools. Kentucky includes certificates of completion awarded to students with disabilities in the total number of high school graduates. NCLB, however, does not permit states to count students with certificates of completion as graduates when calculating the graduation rate. Kentucky, with 19 other states, has chosen to use the National Center for Education Statistics completion rate as the state graduation formula.⁴ The U.S. Department of Education has accepted Kentucky's usage of this formula. To comply with NCLB, Kentucky intends to exclude the number of students with certificates of completion from the number of graduates.

Failing Schools and Districts

Each school district must review schools' performance annually to determine if adequate yearly progress has been made.

No Child Left Behind requires each local school district to use the state assessments and indicators to review the progress of each school annually to determine whether schools are making adequate yearly progress [§1116(a)]. The districts must take appropriate action to improve schools failing to make adequate progress and must also publicize and disseminate the results of the annual review to parents, teachers, principals, schools, and the community. KDE proposes to review schools' performance annually to determine if they have made adequate yearly progress and administer the appropriate, federally required consequences. Biennial reviews and appropriate state consequences will continue with CATS.

Some of the requirements to improve failing schools under CATS and failing schools under NCLB are similar, but they are not identical. Under CATS, schools that score below a certain level are classified as needs assistance and undergo a review and audit process that includes some consequences similar to those provided under NCLB.

⁴ The formula is provided on page 25 of this report.

School Improvement

NCLB requires that a school that fails to make adequate yearly progress for two years be identified for improvement. The district must allow students to attend another school in the district that is not failing. CATS provides for school choice after four years of failure.

If the annual review indicates that a school has failed to make adequate yearly progress in any content area for two consecutive years, the school district must identify that school for improvement [§1116(b)]. The identification of the school must take place before the beginning of the school year following the failure to make adequate yearly progress.

If a school has been identified for improvement, NCLB imposes three requirements the district and the school must meet. First, no later than the first day of the school year following the identification, the school district must provide the school's students with the option to transfer to another public school served by the same district that has not been identified for improvement. The school district must give priority to the lowest achieving children from low-income families. Second, the school must develop or revise a two-year school plan to address the issues that caused it to be identified for improvement. Third, the school district must also provide technical assistance to the school as it develops and implements its plan. If a school fails to make adequate yearly progress for a third year, the district must make supplemental educational services available.

Under CATS, if a school's assessment results place it in the bottom third of schools classified as needing assistance for two consecutive biennia, or four years, that school is required to offer school choice—the option for students to transfer to another public school. KDE intends to amend Kentucky's current regulations to incorporate the additional NCLB requirements of school choice after two years of failure under NCLB.

Corrective Action

Under NCLB, after four years of failure to make adequate yearly progress, the school shall be identified for corrective action.

No Child Left Behind provides that if a school identified for improvement fails to make adequate progress for a fourth consecutive school year, the school district shall identify the school for corrective action [§1116(b)(7)]. The school district shall continue to offer public school choice, provide technical assistance, and make available supplemental educational services. In addition, the school district is required to take at least one of the following corrective actions:

- Replace school staff relevant to the failure;
- Implement a new curriculum;
- Decrease authority of school management;

- Appoint an outside expert to advise the school;
- Extend the school year or school day; or
- Restructure the school's internal organizational structure.

There are no equivalent requirements in CATS for many of these provisions. Some schools classified as needing assistance under CATS do undergo a review and audit process conducted by outside auditors. That process can include personnel evaluations and recommendations regarding retention, dismissal, or transfer of personnel. KDE proposes to amend the current regulations to add the NCLB requirements for schools that fail to make adequate yearly progress for four consecutive years.

School Restructuring

After five years of failure to make adequate yearly progress, the school must be restructured.

NCLB requires that if a school has undergone corrective action for one full school year and fails to make adequate yearly progress for a fifth consecutive year, the school district must prepare a plan to restructure the school [§1116(b)(8)]. The district must implement one of the following alternative governance arrangements for the following school year:

- Reopening the school as a public charter school;
- Replacing all or most of the school staff;
- Contracting with an entity with a demonstrated record of effectiveness to operate the school;
- Turning over the operation of the school to the state educational agency; or
- Restructuring in any other way that makes fundamental reforms.

The district must also continue to offer school choice and make supplemental educational services available. KDE proposes to amend Kentucky's regulations and make legislative recommendations as necessary to include these NCLB requirements.

NCLB requires notification to parents of students attending failing schools before the beginning of the next school year. Under CATS, notification is provided after the beginning of the next school year.

NCLB requires school districts to provide prompt, clear notices to the parents of each student enrolled in a school identified for school improvement, for corrective action, or for restructuring. The notice must explain the meaning of the identification and the reasons for it. NCLB requires decisions on adequate yearly progress to be made and notices to be issued to parents prior to the beginning of the next school year.

Because of the time required to score the open-response and on-demand items included in the CATS assessments, the results are

not available until the fall of the following school year. Obtaining the results earlier could require moving the assessment dates earlier in the school year, giving schools less time to cover the core content before assessment. KDE is negotiating with the assessment contractor to resolve this problem. KDE is considering moving up the assessment dates or seeking to obtain the results of the multiple-choice examinations to allow preliminary decisions on adequate yearly progress to be made by August 1. Final decisions on adequate yearly progress would follow in October after all of the results are returned.

Only Eight States Have Timely Identification of Schools in Need of Improvement

Most states must implement more timely identification of schools that need assistance.

To meet No Child Left Behind's school improvement provisions, most states still need to identify schools in need of assistance in a timely fashion and provide scientifically based assistance to such schools. Currently, only eight states provide timely identification and five states provide technical assistance as required. Around half the states have met the other NCLB school improvement provisions. CATS already includes some school improvement provisions similar to those in NCLB.

NCLB Safe Schools

Each state must implement a statewide policy to allow students to attend safe schools. More than 30 states have addressed school safety to comply with NCLB.

No Child Left Behind requires each state to implement a statewide policy to ensure that students may attend safe schools [§9532]. The policy requires that any student attending a persistently dangerous public school or who becomes a victim of a violent criminal offense while on the grounds of a public school will be allowed to transfer to a safe public school within the local school district. Each state, in consultation with local school districts, will determine the definition of "persistently dangerous public school." At its June 2003 meeting, the Kentucky Board of Education adopted unsafe school choice policies that comply with NCLB.

More Than 30 States Are in Compliance With NCLB Safe School Provisions

More than 30 states, including Kentucky, have addressed school safety to comply with NCLB. Thirty-two states have established criteria for unsafe schools and have enacted a transfer policy for students in unsafe schools. Thirty-one states have established a transfer policy for victims of violent crimes.

Teacher and Paraprofessional Qualifications

NCLB requires all teachers teaching core academic subjects to be highly qualified by the 2005-2006 school year. NCLB defines “highly-qualified” differently according to a teacher’s experience and the grade level taught.

No Child Left Behind requires each school district to ensure that all new teachers hired after the first day of the 2002-2003 school year are “highly qualified” as defined by the Act [§1119]. It also requires that each state develop a plan to ensure that all other teachers who teach core academic subjects are highly qualified no later than the 2005-2006 school year. The term “core academic subjects” refers to English, reading, math, science, foreign languages, civics and government, economics, arts, history, and geography [§9101(11)]. The state’s plan must include an annual increase in the percentage of highly qualified teachers.

The Act defines the term highly qualified somewhat differently according to a teacher’s experience and the grade level being taught [§9101(23)]. A new elementary teacher is highly qualified if the teacher holds at least a bachelor’s degree and has demonstrated subject knowledge and teaching skills by passing a rigorous state test. A current elementary or secondary school teacher is highly qualified if the teacher meets the same requirements, and is certified and licensed to teach in the state (so long as no certification or licensure requirements were waived on any emergency, temporary, or provisional basis). A new middle or secondary school teacher is highly qualified if the teacher holds at least a bachelor’s degree and has passed a rigorous state subject test or completed an academic major, a graduate degree, equivalent coursework, or advanced certification in each of the subjects the teacher teaches.

Kentucky teachers who are emergency certified would not meet requirements to be highly qualified. NCLB requires all paraprofessionals to meet qualification requirements by January 2006. Kentucky’s paraprofessional requirements are less stringent.

Kentucky’s teacher certification requirements are similar to NCLB’s teacher qualification provisions. Kentucky’s regulations do provide specifically for the waiver of some requirements and the issuance of emergency certifications under limited circumstances. Kentucky teachers with emergency certifications would not be considered highly qualified under NCLB.

NCLB also creates new requirements regarding the training and qualifications of paraprofessionals, such as teachers’ aides, who work in public schools [§1119]. The Act stipulates that all new paraprofessionals hired after January 8, 2002, must have either 1) obtained an associate’s degree or higher; 2) completed at least two years of college; or 3) have “demonstrated through a formal assessment they have the knowledge of and ability to assist in instructing reading, writing, and math.” All paraprofessionals must also have a high school diploma. All existing paraprofessionals, regardless of hiring date, must meet the same requirements no later

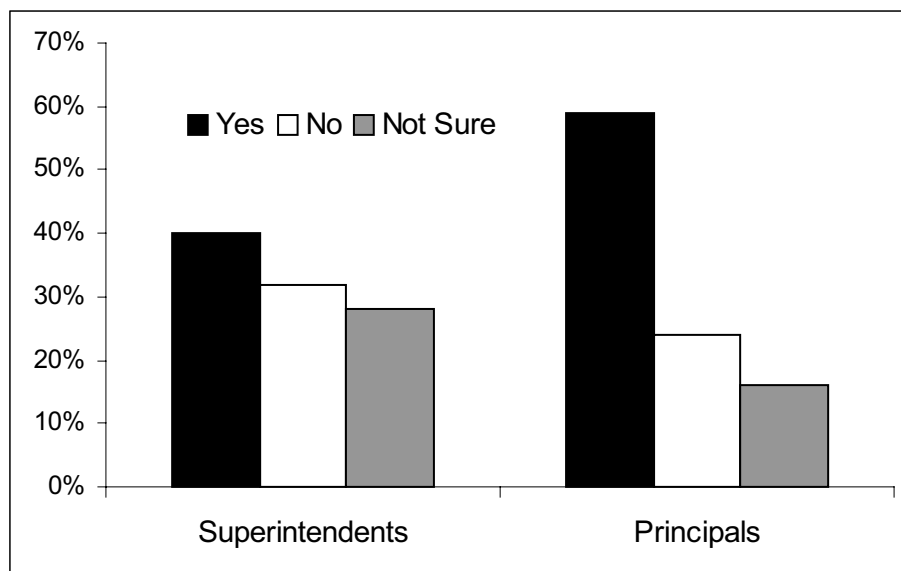
than January 8, 2006. Kentucky statutes only require teachers' aides to have a high school diploma or general equivalency diploma. They do not require teachers' aides to take any formal assessment or obtain additional education.

Kentucky Educators' Opinions on Teacher Qualifications

More than half of Kentucky's principals and about 40 percent of superintendents report that they can have a highly qualified teacher in every classroom by 2005.

Program Review staff elicited opinions from superintendents and principals about the NCLB teacher qualification requirements. Specifically, staff asked whether they could have a highly qualified teacher in every classroom by 2005. As shown in Figure 5.A, more than half the principals and about 40 percent of the superintendents indicated their schools could have a qualified teacher in every classroom by 2005. Around one-fourth of the principals and more than one-third of the superintendents indicated that they would be unable to have a highly qualified teacher in every classroom by 2005.

Figure 5.A
Do You Think Your School/District Can Have a
Highly Qualified Teacher in Every Classroom by 2005?



Source: Program Review staff surveys of superintendents and principals.

Superintendents were also asked why they believed their districts could or could not have a highly qualified teacher in every classroom by 2005. Eighteen percent of responses cited a general lack of highly qualified teachers as a barrier to meeting this requirement, and 16 percent of responses noted it would be very difficult to find highly qualified teachers to teach special

education. Ten percent of responses expressed concern about the difficulty of recruiting highly qualified teachers in rural areas.

**Comments From the CATS Survey:
Highly Qualified Teachers by 2005**

“Some areas where shortages exist make this difficult. For example, in special education we can't be very selective; we have to take whoever we can get. Many are emergency certified.”

“It will be very difficult to replace teachers. In my rural area the teacher shortage is already being felt. There are fewer and fewer applicants to choose from.”

Wisconsin Is the Only State With a Highly Qualified Teacher in Every Classroom

Although most states have tests for new teachers, few states satisfy the other NCLB requirements.

States have made little progress toward satisfying No Child Left Behind's teacher qualification requirements. Although most states have tests for new elementary teachers, very few states satisfy the other NCLB teacher qualification provisions.

Table 5.2 illustrates that most states need to define “highly qualified teacher” and require that teachers have subject matter competence. Two states have implemented highly qualified professional development, and only one state, Wisconsin, has a highly qualified teacher in every classroom.

**Table 5.2
Where States Stand: NCLB Teacher Qualification Provisions***

NCLB Provision	States Meeting Provision	Kentucky (Yes, Partially, No)
Highly Qualified Teacher Definition	10	Partially (State law contains conditional certification provision)
Subject Matter Competence	10	Yes
Test for New Elementary Teachers	41	Yes
Highly Qualified Teacher in Every Classroom	1	No
Highly Qualified Professional Development	2	No

*Includes District of Columbia.

Source: Compiled by Program Review staff from the Education Commission of the States database.

The Estimated Cost of Implementing NCLB

GAO estimated that states could spend between \$1.9 billion and \$5.3 billion on NCLB testing between 2002 and 2008.

The U.S. General Accounting Office (GAO) estimates that states could spend between \$1.9 billion and \$5.3 billion on NCLB testing between 2002 and 2008. GAO's May 2003 report explained that expenditures could vary dramatically based on the types of tests used and how the tests are scored. GAO's major findings were:

- States that use tests with multiple-choice questions, which are machine scored, will expend approximately \$1.9 billion on assessments between 2002 and 2008.
- States that keep the current mix of questions will spend approximately \$3.9 billion on assessments between 2002 and 2008.
- States that use tests with a mixture of multiple-choice and open-response questions that are hand scored will expend approximately \$5.3 billion on assessments between 2002 and 2008.

Federal funding for NCLB assessments is estimated at \$2.7 billion between fiscal years 2002 and 2008. For states to minimize the amount of state dollars they must use to implement No Child Left Behind assessments, GAO recommended that states explore ways to control expenses, such as using computer-administered tests, computerized scoring, or written responses and computer scanning of written responses.

Instead of paying a contractor, Maine, New Hampshire, Rhode Island, and Vermont are working together to create new reading and math assessments.

Instead of paying a contractor to create the required reading and math assessments in grades three to eight, four states have joined forces to create their own assessments to satisfy NCLB. While many states are utilizing off-the-shelf norm-referenced tests, Maine, New Hampshire, Rhode Island, and Vermont are creating their own assessment that will be aligned with all four states' academic standards. They are collaborating to keep costs down and have the new assessments ready to implement by the 2005-2006 school year. The U.S. Department of Education approved the joint effort in January 2003 and awarded the states \$1.78 million to use in creating the assessments.

Based on concerns with the cost of implementing NCLB and the disruption to existing state testing systems, resolutions have been introduced in a few states calling for more federal funding to help implement NCLB. As of July 1, 2003, none of the measures had been enacted.

Works Cited

Chaplin, Duncan, and Christopher B. Swanson. *Counting High School Graduates When Graduates Count: Measuring Graduation Rates Under the High Stakes of NCLB*. Urban Institute: Washington, February 2003.

Education Commission of the States. *No Child Left Behind*. July 29, 2003
<http://nclb.ecs.org/nclb/>

Education Week. *Quality Counts 2003: If I Can't Learn From You..* Volume 22 (17). Bethesda, Maryland, January 9, 2003.

Goertz, Margaret, Mark Duffy, and Kerstin Carlson-Le Floch. *State Assessment and Accountability Systems: 50 State Profiles*. Consortium for Policy Research in Education. March 2001.

Greene, Jay P. *High School Graduation Rates in the United States*. Manhattan Institute: New York, April 2002.

Hambleton, R.K., R.M. Jaeger, D. Koertz, R. Linn, J. Millman, and S.E. Phillips. *Review of the Measurement Quality of the Kentucky Instructional Results Information System, 1991-1994*. A report prepared for the Office of Educational Accountability. Kentucky General Assembly: Frankfort, 1995.

State of Texas. Office of the State Auditor. *An Assessment of the Texas Education Agency's Monitoring Systems for Public Education*. Austin, August 1996.

Sum, Andrew, and Paul Harrington. *The Hidden Crisis in the High School Dropout Problems of Young Adults in the U.S.: Recent Trends in Overall School Dropout Rates and Gender Differences in Dropout Behavior*. Center for Labor Market Studies: Boston, February 2003.

United States. General Accounting Office. *Title I: Characteristics of Tests Will Influence Expenses; Information Sharing May Help States Realize Efficiencies*. GAO-03-389. Washington, May 2003.

University of the State of New York. State Education Department, Office of Audit Services. August 4, 2003 <http://www.oms.nysed.gov/oas/Audit%20reports/SchoolDistricts&EPE/SchoolDistricts&EPE.htm>

Appendix A

Summary of Statutes and Regulations Governing CATS

Statutes

Legislative Intent

KRS 158.645 expresses recognition that public education involves responsibilities shared among the local communities, parents, students, school employees, and state government, and that the cooperation of all is needed to achieve desired outcomes.

This statute is also a declaration of legislative intent to create a system of public education that will allow and assist all students to acquire the capacities below:

1. Communication skills;
2. Knowledge to make economic, social and political choices;
3. Core values and qualities of good character;
4. Understanding of governmental processes;
5. Sufficient knowledge of self and mental and physical wellness;
6. Sufficient grounding in the arts to appreciate one's cultural and historical heritage;
7. Sufficient preparation to intelligently choose and pursue one's life's work;
8. Skills to enable one to compete favorably with students in other states.

Goals for Schools

KRS 158.6451 is a declaration of goals for the Commonwealth's schools.

Schools shall expect a high level of achievement of all students and shall develop their students' ability to:

1. Use basic communication and math skills;
2. Apply core concepts and principles learned in classes to real-life situations;
3. Become self-sufficient individuals of good character;
4. Become responsible members of a family, work group, or community;
5. Think and solve problems in a variety of situations;
6. Connect and integrate experiences and new knowledge from all subject matter fields.

Schools shall also increase their students' attendance rate, and reduce their dropout and retention rates. Schools shall be measured on the proportion of students to make a successful transition to work, post-secondary education, and the military.

This statute further requires the Kentucky Board of Education (KBE) to disseminate to local districts a model curriculum framework directly tied to the goals, outcomes, and assessment strategies set forth in the statutes. The framework shall provide direction to local districts and schools as they develop their curriculum. The framework shall identify teaching and assessment strategies, instructional material resources, ideas regarding incorporating the resources of the community, a directory of model teaching sites, alternative ways of using school time, and strategies to incorporate character education throughout the curriculum.

School Curriculum, Assessment,
and Accountability Council

KRS 158.6452 creates the School Curriculum, Assessment, and Accountability Council to study, review and make recommendations regarding setting academic standards, assessing learning, holding schools accountable and assisting them to improve. The Council shall advise the KBE and the LRC on issues related to the development and communication of academic expectations and core content for assessment, the development and implementation of the assessment and accountability program, the distribution of rewards and imposition of sanctions, and assistance for schools to improve their performance under KRS 158.6453, 158.6455, 158.782, and 158.805.

The composition of the Council shall be seventeen voting members appointed by the Governor, drawn from various defined classifications, with broad geographical representation of all school levels, equal representation of both sexes, and minority representation that reflects the composition of the Commonwealth. The Council is attached to the Department of Education (KDE) for administrative purposes.

CATS and Its Components

KRS 158.6453 defines and creates the Commonwealth Accountability Testing System (CATS) to ensure school accountability for student achievement of goals set forth in KRS 158.645 and KRS 158.6451 and places the responsibility for the creation and implementation of CATS with the KBE. The assessment program shall not measure students' ability to become a self-sufficient individual or to become a responsible member of a family, work group or community.

CATS shall include the following components:

1. Valid and reliable, customized, or commercially-available norm-referenced test that measures the core content for assessment;
2. Questions to assess student skills in reading, mathematics, science, social studies, the arts, the humanities, and practical living and vocational studies, and an on-demand assessment of student writing;
3. Writing portfolios consisting of samples of student work;
4. Performance assessment events for schools with students enrolled in performing arts organizations;
5. Technically sound longitudinal comparison of the assessment results for the same students.

Role of Teachers in
Designing CATS

Kentucky teachers shall have a significant role in designing CATS. CATS shall be designed to:

1. Measure grade appropriate core academic content, basic skills, and higher-order thinking skills and their application;
2. Provide valid and reliable scores;
3. Minimize time spent by teachers and students on assessment.

CATS results must be reported to districts and schools no later than 150 days following the first day the assessment can be administered.

Establishing Validity of CATS

The KDE must gather information to establish the validity of CATS and must develop a biennial plan for validation studies. The statute sets forth the minimum studies that must be included in the plan.

In addition to creating and implementing CATS, the KBE also has responsibility for assisting local school districts and schools in developing and using continuous assessment strategies needed to assure student progress.

School Report Cards

The KBE is also required to promulgate regulations to establish the components of a school report card that clearly communicates with parents and the public

about school performance. The report card shall be sent to the parents of the students of the districts, and a summary of the results for the district shall be published in the newspaper with the largest circulation in the county. The report card is required to include the following information, at a minimum, reported by race, gender, and disability when appropriate:

1. Student academic achievement, including the results from each of the assessments administered pursuant to this statute;
2. Nonacademic achievement, including the school's attendance, retention, dropout rates, and student transition to adult life;
3. School learning environment, including measures of parental involvement.

National Technical
Advisory Panel on
Assessment and
Accountability

KRS 158.6454 provides for the appointment of a National Technical Advisory Panel on Assessment and Accountability (NTAPAA) consisting of no fewer than three professionals with a variety of expertise in education testing and measurement. The panel shall provide advice regarding the implementation of KRS 158.6453 and 158.6455.

Legislative Intent for School
Accountability System

KRS 158.6455 is a declaration of legislative intent and provides the framework for the school accountability system.

It is the intent of the General Assembly that *"schools succeed with all students and receive the appropriate consequences in proportion to that success."*

School Accountability
Index

The KBE is given responsibility for promulgating administrative regulations to establish a system for identifying and rewarding successful schools. Rewards are to be distributed to successful schools based on the number of certified staff employed in the school on the last working day of the year of the reward. The KBE shall identify reports, paperwork requirements, and administrative regulations from which high-performing schools shall be exempt. Effective July 1, 2006, the KBE shall reward schools that exceed their improvement goal and have an annual average dropout rate below 5%.

The KBE is given responsibility for promulgating regulations which set forth the formula for a school accountability index to classify schools every two years based on whether they have met their threshold level for school improvement. The formula must reflect the academic goals set forth in KRS 158.6451.

The KBE is further required to promulgate regulations to establish appropriate consequences for schools failing to meet their threshold. The consequences shall be designed to improve teaching and learning and may include, but not be limited to:

1. A scholastic audit process to determine the appropriateness of a school's classification and to recommend needed assistance;
2. School improvement plans;
3. Eligibility to receive Commonwealth school improvement funds under KRS 158.805;
4. Education assistance from highly skilled certified staff under KRS 158.782;
5. Evaluation of school personnel; and
6. Student transfer to successful schools.

Scholastic Audits

The KBE is also required to promulgate regulations establishing guidelines for conducting scholastic audits, which shall include the process for:

1. Appointing and training audit team members;

2. Reviewing a school's learning environment and efficiency, students' academic performance, and the quality of the school council's data analysis and planning;
3. Evaluating each certified staff member;
4. Making a recommendation to the KBE about the appropriateness of the school's classification and a recommendation concerning the assistance required by the school to improve teaching and learning.

For information purposes, the KBE is also required to conduct scholastic audits in a sample of schools that achieved their goals and report the findings to the public.

The statute further requires the KBE to promulgate regulations which establish a formula for school accountability and a school improvement goal for each school for the 1998-1999 and 1999-2000 school years. Rewards and audits were to be assessed for those years as set forth in the statute.

The KBE was given discretion to promulgate regulations which created a system of district accountability similar to the system of school accountability, including a formula for accountability, goals for improvement, rewards for leadership in improving teaching and learning, and consequences for a district's failure to achieve its goals.

Appeals Process

The KBE is also required to promulgate regulations establishing a process by which a school can appeal a performance judgment it considers grossly unfair. KBE may adjust a performance judgment when evidence warrants the conclusion that the judgment is based on fraud or a mistake in computations, is arbitrary, is lacking any reasonable basis, or if there are significant new circumstances occurring during the biennial assessment period which are beyond the control of the school.

Definitions

KRS 158.6457 provides the following definitions:

1. "Accountability index" means the statistic that combines a school's academic and nonacademic factors;
2. "Core content for assessment" means the content identified for all students to know that is to be included on the state assessment; and
3. "Nonacademic factors" means the statistic that describes school success on increasing attendance and decreasing retention and dropout rates.

KDE to Develop Plan to Implement CATS

KRS 158.6458 requires the KDE to develop a plan to implement CATS and to provide quarterly reports to the Interim Joint Committee on Education on its progress in the following nine areas:

1. Establishing a consistent structure of test components, test distribution, and test administration procedures;
2. Beginning a new cycle of equating procedures and conducting appropriate equating analyses;
3. Publishing informative guides for interpreting school accountability index score changes;
4. Reviewing school accountability classifications to assure their construct validity;
5. Maintaining and strengthening the annual audit of portfolio scores;
6. Developing and implementing a validity research plan;
7. Establishing additional routine audits of key processes;
8. Maintaining a library of technical documents and producing an annual technical report; and

9. Maintaining a vigorous ongoing program of research and documentation.

Kentucky Institute for
Education Research Board

KRS 158.646 creates the Kentucky Institute for Education Research Board (KIERB) and defines its purpose, duties and membership. The two-fold purpose of the KIERB is:

1. Solicit and raise funds to support the independent evaluation of the KERA and related activities;
2. Serve as a stimulus and clearinghouse for KERA-related research projects.

The duties of KIERB are:

1. Cause an in-depth evaluation of the impact of KERA to be performed;
2. Make recommendations regarding enhancement of benefits of KERA and expansion and improvement of services to students;
3. Develop the capacity to manage and coordinate research, conduct research, and design and implement a comprehensive educational data information system; and
4. Prepare an annual report of its activities.

KIERB is to consist of ten members initially appointed by the Governor. The members are required to enact by-laws governing membership making the board self-perpetuating.

Education Assessment and
Accountability
Subcommittee

KRS 158.647 creates the permanent subcommittee, Education Assessment and Accountability Review Subcommittee (EAARS) and defines its membership and purpose.

It is to be composed of eight members, four from each chamber. This subcommittee has responsibility for reviewing administrative regulations and advising KBE concerning the implementation of CATS, and for advising and monitoring the Office of Education Accountability.

Achievement Gaps

KRS 158.649 defines the term “achievement gap” and requires the KDE to provide each school with performance data on its students disaggregated by race, gender, disability, English proficiency, and participation in the free and reduced price lunch program.

The statute requires each local board of education to adopt a policy for reviewing the academic performance for various groups of students. It further requires the local boards to establish biennial targets no later than December 1, 2002, for each school for reducing identified achievement gaps.

By February 1, 2003, and each February 1 in odd-numbered years thereafter, the school-based decision making council shall set the school’s biennial targets for eliminating achievement gaps and submit them to the superintendent for consideration. With the agreement of the superintendent, the targets are submitted to the local board of education for adoption.

By April 1, 2003, and each April 1 in odd-numbered years thereafter, the school council, with the involvement of parents, faculty, and staff, shall review the data and revise the consolidated plan to include the biennial targets, strategies, activities, and a time schedule calculated to eliminate the achievement gap. The principal shall convene a public meeting at the school to present and discuss the plan prior to its submission to the superintendent and local board of education. At a minimum, the plan shall address the following areas:

1. Curriculum alignment;
2. Evaluation and assessment strategies to monitor and modify instruction to meet student needs and support proficient student work;
3. Professional development;
4. Parental communication and involvement;
5. Attendance improvement and dropout prevention; and
6. Technical assistance.

Local boards shall determine if each school has achieved its biennial targets for each group of students. If a school has not met its target, the board shall require the council to submit its revisions to the consolidated plan describing the use of funds to reduce the school's achievement gap. The plan must address how the school will meet the academic needs of the group of students at issue.

If a school fails to meet its targets to reduce the achievement gap for any student group for two successive biennia, the superintendent shall report that failure to the commissioner for education. The school's consolidated plan shall be subject to review and approval by the KDE and the school shall submit an annual status report. If a school meets its biennial target for reducing the achievement gap, the council is no longer required to seek approval of its plan.

Regulations

Writing Portfolio Procedures

703 KAR 5:010 establishes procedures to reduce the teacher and student time involved in preparing a writing portfolio.

A five-piece portfolio shall be produced in 12th and 7th grades and a four-piece portfolio shall be produced in 4th grade. Schools and districts shall also develop a procedure to collect writing pieces at nonaccountability levels that are appropriate types of writing for portfolio categories. These pieces may serve as rough drafts that can be edited for inclusion in the accountability portfolio or they may be included as finished products.

Each school and district shall provide support for teachers to attend professional development focused on the types of writing assessed in the portfolio. Each school and district shall develop procedures for scoring portfolios that include an adequate number of teacher scorers to limit the number of portfolios scored by any one teacher to thirty. Conferencing on portfolio pieces shall be focused on one or two areas of need, addressing patterns of errors that occur frequently.

Teachers shall allow students to use word processing during the development of writing pieces or allow students to submit pieces in their own handwriting. Teacher assigned writing tasks shall relate to the content being studied. Teachers shall choose content area readings that represent the kind of writings the students are asked to include in the portfolio, allowing the discussion of content and writing form to occur at the same time.

Scoring System

703 KAR 5:020 establishes the scoring system for determining successful schools, school rewards, and classifications of schools within the school accountability program.

Section 1. The following terms are defined:

1. Academic index;
2. Accountability index;
3. Accountability level;

4. Assistance line;
5. Alternate portfolio;
6. Alternate portfolio scores;
7. Baseline accountability index;
8. Gained population;
9. Goal line;
10. Growth accountability index;
11. Lost population;
12. Nonacademic index;
13. Reward share;
14. School classification;
15. School;
16. School recognition points;
17. Stable population;
18. Standard error of measurement;
19. Standing of a school;
20. State goal;
21. Student achievement levels;
22. Target biennium;
23. Threshold; and
24. Writing portfolio scores.

Section 2. This section assigns the points awarded to students at each achievement level for purposes of calculating the academic indices as follows:

1. Nonperformance – 0 points;
2. Medium novice – 13 points;
3. High novice – 26 points;
4. Low apprentice – 40 points;
5. Medium apprentice – 60 points;
6. High apprentice – 80 points;
7. Proficient – 100 points; and
8. Distinguished – 140 points.

Values for attendance rate and successful transition to adult life rate shall be the actual percentage reported. The values for retention rate and dropout rate shall be 100 minus the actual percentage reported. Alternate portfolio scores shall be included so that they contribute the same weight as do scores for other students.

Section 3. This section defines the two components of the accountability index and explains how to calculate those components. It assigns various weights by content area and grade level which are to be used in calculating component one. Component two is to be calculated using a national norm-referenced test and the scores assigned by achievement level set forth above.

Section 4. This section addresses how scoring should be undertaken for those schools that do not conform to standard grade configurations. Schools may request a waiver from scoring requirements and specify other combinations of schools and assessment data so long as all students in an accountability grade are included.

Section 5. This section provides that if a school has more than one accountability level (elementary, middle, or high school), that school's accountability index shall be the average of the academic and nonacademic data for the school.

Section 6. This section defines reconfigured schools and provides for their treatment in accountability decisions. Reconfigured schools are those in which

less than 80% of the school's population is stable. A reconfigured school shall have the performance judgment that would have applied to the district at that level. In the alternative, a school district may submit a plan for reconstituting baseline data taking into consideration the changes in service areas and assuring that local district calculations are accurate and include all student data.

Section 7. This section provides how to establish expected levels of growth for each school and establishes five points of school recognition for the purpose of recognizing school standing.

School Classifications and
Rewards

Section 8. This section defines the following school classifications:

1. Meets goal;
2. Progressing;
3. In need of assistance; and
4. Commonwealth pace-setter.

A school's classification is determined through a comparison of the school's growth accountability index for a biennium with a corresponding goal point and assistance point. This section also sets forth requirements schools must meet in order to receive rewards. A pace-setter school shall be entitled to one share of rewards if not otherwise receiving rewards for growth.

Section 9. This section establishes the levels of rewards available to schools. A school classified as "meets goals" shall earn three shares of rewards. A "progressing" school shall earn one-half share of rewards. Additionally, schools that meet or exceed school recognition points are entitled to a one-time reward of one share for each point met or exceeded. Numbers of shares earned are multiplied by the number of certified staff to determine the final reward amount.

School Categories and
Accountability

703 KAR 5:040 defines the characteristics of A1-A6 schools and explains how accountability indices and rewards will be determined for the different categories of schools.

The school categories are defined as follows:

1. An "A1" school means a school under administrative control of a principal or head teacher and which is eligible to establish a school-based decision making council. An A1 school is not a program operated by, or as a part of, another school;
2. An "A2" school means a district-operated, totally vocational-technical school;
3. An "A3" school means a district-operated, totally special education school;
4. An "A4" school means a district-operated, totally preschool program;
5. An "A5" school means an alternative school designed to provide services to at-risk populations with unique needs;
6. An "A6" school means a district-operated instructional program in a nondistrict-operated institution or school.

Accountability indices and related statistics shall be calculated only for those schools classified as A1 schools. Any A2-A6 school from which graduates receive a high school diploma or a certificate shall determine and report data regarding the graduates' transition to adult life and shall indicate the A1 sending school to which the data should be attributed.

Nonacademic data collection procedures shall apply to all A1 – A6 schools.

For the purposes of rewards and assistance, staff of the A2 – A6 schools shall be viewed as providing a service to the total district and shall receive rewards or assistance as appropriate for the district as a whole.

An A2 – A6 school serving multiple districts shall be eligible for rewards if more than 10% of its total membership is generated from a school which has qualified for rewards. An A2 – A6 school shall be subject to assistance if more than 50% of the aggregate membership of the school is generated from the district being required to receive assistance.

Appeals of Performance
Judgments

703 KAR 5:050 establishes procedures for a school to appeal a performance judgment it considers to be grossly unfair.

It defines the following terms:

1. Baseline accountability index;
2. Growth accountability index; and
3. Performance judgment.

A school may request a data review by submitting the request to the Commissioner of Education (COE) within fourteen days after the KDE officially releases the performance judgments to the public. If the performance judgment is revised or the school is not satisfied with the results of the data review the school shall submit a written appeal of the performance judgment to the COE within 30 days after the school has received notification of the review results.

A school may appeal a performance judgment by submitting a written appeal to the COE within forty-five days after the Department of Education officially releases the performance judgments to the public. The appeal must clearly identify the basis for the wrongful effect on the accountability indice(s) and shall detail the requested adjustment to be made to one or more of these indices.

The COE shall appoint a committee to review the pending appeals and make recommendations to the COE as to whether to dispute an appeal. A hearing officer shall conduct a hearing and shall submit a written recommended order to the KBE for the board's consideration in rendering its final order.

Special Student Populations

703 KAR 5:070 establishes procedures for the inclusion of special student populations in the state-required assessment and accountability programs.

It incorporates by reference the 27-page document entitled, "Inclusion of Special Populations in the State-Required Assessment and Accountability Programs."

**Inclusion of Special Populations in the State-Required
Assessment and Accountability Programs**

Special populations include:

1. Students with disabilities;
2. Students who attend schools classified as A2 – A6 schools and state agency children;
3. Students whose primary language is not English;
4. Students receiving instruction in home/hospital settings (i.e. home-bound instruction, not home schools); and
5. Students who have temporary medical conditions that necessitate accommodations or modifications or both for participation.

The provisions of this regulation apply to all students except those 21 years of age or older who are part-time students attending less than 6 hours per day and those enrolled in the Adult General Education Diploma Program (GED).

All students with disabilities shall participate in the state-required Assessment and Accountability Programs. A small percentage of students shall participate in the Alternate Portfolio Assessment Program.

Students with limited English proficiency shall be included if they have been in the same school or district for one full academic year prior to the year of the assessment in question, or in an English-speaking school for two full academic years prior to the year of assessment in question.

Students receiving instruction in home/hospital settings shall participate in the state-required Assessment and Accountability Programs unless a school or district has obtained a signed Physician's Certificate of Student Exemption describing the medical condition that warrants exempting a student from all or portions of the assessments.

Students with disabilities are students who meet the criteria under KRS 157.200 and the Kentucky Administrative Regulations related to Exceptional Children or Section 504 of the Rehabilitation Act of 1973.

For students with disabilities, the Admissions and Release Committee (ARC) or Individual Education Program (IEP) or 504 Committee shall determine on an individual basis how the student will be included in the state-required Assessment and Accountability Programs.

Students with Disabilities

Section 1 – Inclusion of Students with Disabilities

This section of the regulation sets out three options for inclusion and describes which students shall participate under each option:

Participation with no accommodations or modifications.

This option includes students who have a remedial plan but who have not been identified as disabled, students referred for evaluation but for whom the evaluation process has not been completed, and those students with disabilities who are not receiving special education and related services.

Participation with accommodations or modifications or both.

This option includes students who meet the eligibility criteria for one of the disability categories under the Kentucky Administrative Regulations, have a current IEP, and are receiving specially designed instruction. The students must also currently be using accommodations or modifications as part of his or her regular instructional routine. The accommodations or modifications must be related to the individual student's needs and the impact of the disability and must be specified in the student's IEP.

Students who meet the eligibility criteria in Section 504 of the Rehabilitation Act of 1973 for having a physical or mental disability which substantially limits one or more major life activities and have a current 504 Plan may use accommodations for the assessment if:

1. The accommodations or modifications are part of the student's regular instructional routine;

2. The accommodations or modifications are related to the individual student's needs and the impact of the disability; and
3. The accommodations or modifications are specified in the student's 504 Plan.

Alternate Portfolio
Assessment Program

Participation in the Alternate Portfolio Assessment Program

This option includes only those students who meet all the criteria for the certificate program as stated in 707 KAR Chapter 1 related to Exceptional Children and the Program of Studies (704 KAR 3:303).

The results of the Alternate Portfolio Assessment shall count in the accountability calculations and shall be equivalent to the impact of a student participating in the regular Assessment and Accountability Programs process.

Alternate portfolios shall be completed once each in elementary, middle, and high school. The elementary portfolio shall be completed in the fourth grade, the middle school portfolio in the eighth grade and the high school portfolio in the last full year of school.

Students with disabilities in non-graded programs who do not participate in the Alternate Portfolio Assessment Program shall be required to participate in the state-required Assessment and Accountability Programs at their present appropriate grade assignment.

A student with disabilities who turns 21 years old during a school year may "age out" of school without completing the school year and participating in the assessment program. If this is a possibility, the student shall be included in the assessment during the school year prior to turning 21.

Students who skip a grade within the accountability system shall still participate in the assessment components associated with the grade being passed through. With the exception of students who participate in the Alternate Portfolio Assessment Program, schools who serve disabled students shall be held accountable for these students' transition to adult life using the same standards applied for all other students.

Students who participate in the Alternate Portfolio Assessment Program may be considered to have made a successful transition to adult life if they:

1. Make a successful transition as defined for the regular population;
2. Enroll as a full or part-time student at a post-secondary vocational school or adult education program preparing students for integrated work;
3. Work in an integrated setting at least 10 hours per week;
4. Participate in supported employment; or
5. Transition to community rehabilitation and their training or employment takes place in an integrated environment.

Students in A2 – A6 Schools

Section 2 – Inclusion of Students in A2 – A6 Schools and State Agency Children

Students attending schools classified as A2 – A6 shall be included in the overall accountability program. These students' scores shall be tracked back to the A1 schools that would have served them had they not required the services offered by the A2 – A6 schools they attend.

Students with Limited
English Proficiency

State agency children shall have the same assessments administered and the results shall be included in the accountability index of the last A1 school the student attended or the school the student would have attended in that district.

Section 3 – Inclusion of Students Whose Primary Language is Not English

Schools shall assess and be held accountable for all students with limited English proficiency who have been in the same school or district for one full academic year prior to the year of assessment or in an English-speaking school for two full academic years. School personnel shall determine whether students who have been enrolled less time shall participate in the assessment programs.

School personnel shall determine on an individual basis whether students with limited English proficiency will participate in state-required assessments with or without accommodations or modifications or both. A student with limited English proficiency may use accommodations or modifications or both if the student:

1. Meets the criteria as a student with limited English proficiency;
2. Has evaluation data that demonstrates a need for accommodations or modifications;
3. Has a current Program Services Plan that includes accommodations or modifications; and
4. Is participating in instructional programs and services to meet the student's language and academic content needs.

Accommodations or modifications shall be related to the individual student's needs and shall be made in the best interest of the student. Specific documentation must be provided if accommodations or modifications are needed.

Accommodations or modifications may include various administration strategies such as paraphrasing instructions or translating the text of questions. They may also include various response strategies such as allowing students access to a foreign language dictionary and use of grammar and spell-check systems.

A student with limited English proficiency who has been in the same district for less than one full academic year prior to the year the portfolio is due, or in an English speaking school for fewer than two years preceding, may be exempted from the portfolio assessment based on an English language proficiency assessment, the student's Program Services Plan, and the normal ongoing delivery of instructional services.

A student with limited English proficiency may submit a portfolio in a language other than English if:

1. The student's daily instruction and class work are conducted in the student's native language; and
2. The local scorer or a scorer hired by the district is both fluent in that language and trained to score the portfolio.

Home and Hospital
Students

Section 4 – Inclusion of Students Receiving Instruction in Home/Hospital Settings

School personnel shall determine on an individual basis how each such student will participate in the state-required assessment programs. A student shall either participate fully or is exempted from the state-required test upon verification by a physician of an illness or injury that prohibits the student from participating in one or more assessment components. If a student with disabilities is receiving

instruction temporarily or long-term in a home/hospital setting, the procedures described in Section 1 shall be followed.

Students with Temporary
Medical Conditions

Section 5 – Inclusion of Students with Temporary Medical Conditions That Necessitate Accommodations or Modifications or Both for Participation

Students who become injured or develop an ailment before or during the testing window may be allowed appropriate accommodations or modifications or both to allow their participation in the state-required assessment programs.

Appropriate
Accommodations and
Modifications

Section 6 – Conditions for Implementing Accommodations or Modifications or Both

Accommodations or modifications shall be age-appropriate and related to both the student's verified disability and specially-designed instruction. They shall be part of the student's ongoing instructional program and not introduced for the first time during assessment. They shall not inappropriately impact the content being measured.

There are a variety of accommodations and modifications that may be appropriately used for students with disabilities, including:

1. Readers;
2. Scribes;
3. Paraphrasing;
4. Use of technology and special equipment;
5. Extended time;
6. Reinforcement and behavioral modification;
7. Manipulatives;
8. Prompting or cueing; and
9. Interpreters.

Any individual who provides assistance to a student with disabilities during the assessment shall be trained in his role and shall abide by confidentiality laws, ethics provisions, and the conditions for use as described in the student's IEP. Particular forms of the test should be randomly distributed just as with students without disabilities, with the exception of students with limited English proficiency as described in Section 3.

Readers. If listening to a reader is the normal mode through which the student is presented regular print materials, reading assessments may be read to a student. The ARC or 504 Committee shall have considered under what conditions a student will use a reader on a routine basis during instruction. On-demand tasks may be read to students under certain, specified conditions. A reader shall not inappropriately impact the content being measured. A reader shall read information as written and shall not use information to lead the student to specific information needed to answer questions and shall not point out parts of a task or question skipped by the student. A reader shall re-read portions only if specifically requested by the student.

Scribes. The ARC or 504 Committee shall consider under what conditions a student will use a scribe on a routine basis during instruction. A scribe may be used for state assessments under certain, specified conditions. Technology and natural supports shall be used prior to the more intrusive process of using a scribe. A scribe shall not inappropriately impact the content being measured. A scribe's role shall be to record the student's work to allow the student to reflect

what the student knows and is able to do while providing the student with an alternative means to express his thoughts and knowledge.

Paraphrasing. The ARC or 504 Committee shall consider under what conditions a student will use paraphrasing on a routine basis during instruction. On-demand tasks may be paraphrased under specified conditions. Paraphrasing for the state assessment shall be consistent with classroom instruction and includes repeating or rephrasing the directions, prompt, or situation, but shall not include defining words or concepts or telling a student what to do and the order in which it should be done. Stories and content passages may not be paraphrased. A paraphraser shall not inappropriately impact the content being measured.

Use of Technology and Special Equipment. The ARC or 504 Committee shall consider under what conditions a student may use technology on a routine basis during instruction. During the state-required assessment, a student with a disability may use special equipment that is part of the student's regular instructional routine. If it is necessary for a student with special needs to complete written work on a computer and this procedure is routinely used in the student's regular instructional program and noted in his or her IEP or 504 Plan, it may be used when responding to open-response questions while participating in the state-required assessment under specified conditions intended to protect the security, confidentiality and integrity of the assessment.

Extended Time. Students with disabilities shall be allowed extended time to complete items on state-required tests if they have IEPs or 504 plans that stipulate extra time is needed and if extended time is an accommodation for assessments and completion of assignments as part of their daily instructional routine.

Reinforcement and Behavior Modification Strategies. Students with disabilities who have IEPs or 504 plans that stipulate the use of reinforcement or behavior modification strategies and for whom the use of such strategies are implemented during routine instruction, may use these strategies on the state-required assessment. They may also be implemented for any student who displays aggressive or disruptive behavior during testing.

Manipulatives. Manipulatives may be used to complete the state-required tests and the development of portfolios if they are a strategy used by the student to solve problems routinely during instruction and the use of manipulatives is described in the student's IEP or 504 Plan.

Prompting or Cueing. The ARC or 504 Committee shall consider under what conditions a student will use prompting and cueing on a routine basis during instruction. If a student uses a cue card or other strategy on a daily basis during instruction as stipulated by the student's IEP or 504 Plan, the student may use the cue card or strategy during the state-required assessment. The teacher shall not draw figures, suggest leading sentences, or provide content during the administration of the state-required standardized assessment.

Interpreters. The state-required tests may be translated to the student in sign language under specified conditions, including the use of sign language in the student's IEP or 504 Plan and the use of signing as part of the student's regular instructional routine. Signing shall not be a replacement for technology or reading instruction. The interpreter shall not indicate correct answers to test items.

“Administration Code for Kentucky’s Educational Assessment Program” is adopted and incorporated by reference.

**Administration Code for Kentucky’s Educational
Assessment Program**

I. Rationale.

This document describes the practices considered appropriate in preparing students for assessments, in administering them, and in providing for proper security of the assessment materials. No test preparation practice shall violate the ethical standards of the education profession in 704 KAR 20:680. No test preparation practice shall increase students’ test scores on the statewide assessment components without simultaneously increasing students’ ability to apply the content tested to real life situations. All assessment work shall be done entirely by the student.

**Appropriate Assessment
Practices**

II. Appropriate Assessment Practices.

Each individual involved in any component of the assessment must read, sign and comply with the Administration Code and receive training on it. Each test administrator or proctor shall sign a verification form stating he or she has received and read this Code and instruction manual.

Test Security. District assessment coordinators, administrators, and teachers shall ensure the security of the assessment materials before, during and after test administration. It is appropriate for teachers to know and teach the concepts measured by the statewide assessment, but secure test materials shall not be reproduced in any way nor shall notes be taken regarding any secure test item. Tests shall be distributed in the order in which they are received in shrink-wrapped packages. No one may have test booklets without authorization. No one may show items in the test booklets to anyone not administering the test. No one may reveal the content of any secure test item or use that knowledge to prepare students for the assessment. Test administrators must destroy any notes, drafts, or scratch paper produced by students and must ensure that any testing materials reused from previous years are free of any marks.

In those instances in which computer technology must be used to provide access to tests, secure materials shall be scanned to and stored on floppy disks that shall be returned to the contractor. When space requirements are prohibitive, the material may be scanned to larger disk drives if the district staff can assure the security of the assessment. Student responses which reveal that the student may cause harm to himself or others or is suffering abuse may be copied in relevant part and turned over to appropriate authorities.

Procedures for Reporting Errors in Assessment Materials. The test item containing the error shall not be reproduced. Rather, the location of the error shall be identified and the error shall be summarized for the District Assessment Coordinator who shall notify KDE.

Classroom Materials. Materials may be placed on classroom walls and bulletin boards for instructional purposes and lesson plans shall indicate the relationship between the materials and instruction. Staff shall not place materials specifically designed for assessment on classroom walls or bulletin boards.

Dictionaries and thesauri may be used only on the writing-on-demand subtest. Students shall have access to calculators as designated in the administration manuals. Blank writing or graph paper and highlighters or markers may be made available. Other information and materials not sent as part of the assessment materials shall not be made available to students. Students shall not leave the testing area to gain access to calculators or other resources. Materials placed on classroom walls for instructional purposes shall not be moved to other locations for assessment purposes.

Administration Practices. Building personnel and District Assessment Coordinators shall schedule test administration; arrange for adequate staff to administer the assessment; prepare an accurate student accountability roster; and ensure that assessment materials are kept secure before, during, and after testing sessions.

During testing, words of encouragement and general instructions that do not imply evaluation of student work are permissible. Accommodations or modifications may be provided if consistent with a student's IEP, 504, or LEP Plan and the routine delivery of instructional services. Tests should be scheduled to avoid conflicts with lunch. Interval or restroom breaks may be conducted but the integrity of testing shall not be affected. Test sections shall be administered in the order in which they appear in the test booklets. Time limits and specific directions in the manuals shall be observed. Test administrators shall circulate throughout the testing site to monitor students as they work. When students need extended time to complete a test session, the additional time shall begin immediately following the initial administration.

During testing, test administrators shall not answer student questions that would aid the student in responding to any item on the test nor shall they assist the student in understanding the question. Test administrators shall not encourage students to edit their responses by providing any evaluation of student work. Students shall not take more than a single school day to complete a testing session except where there is a documented student illness or emergency. A student shall not be left alone in a room to take the test nor shall she be allowed to take a test booklet or answer booklet out of the testing area without supervision.

Disciplinary Practices and Student Motivation. Administrators may direct students to apply themselves to the task at hand but shall not give any direction that would enable a student to better understand the task or to gain advantage in responding to the task. Student responses may be visually scanned after the testing session to determine disciplinary problems. If disciplinary problems are determined to exist, students shall not be allowed to modify their initial response. If a student's responses are found to contain inappropriate language, the student may be instructed to answer the questions again for disciplinary purposes. Both the original and rewritten responses shall be submitted to the testing contractor with the rewritten ones clearly marked NOT TO BE SCORED—ITEMS RETAKEN FOR DISCIPLINARY PURPOSES. Student responses may be visually scanned during or after the testing session to determine good faith efforts but no evaluative statements shall be made until the entire assessment has been administered and submitted to the District Assessment Coordinator.

Writing Portfolios

Writing Portfolios. Each portfolio entry is to be evaluated by certified personnel, trained to apply the same set of standards in the same manner from student-to-student, from school-to-school, and from year-to-year.

Teachers may provide opportunities for writing appropriate for inclusion in the portfolio and may allow ample time for preparation of portfolio entries in the

classroom and may allow student work outside of class. Teachers may ask questions to clarify aspects of the student's work and may indicate where errors occur and ask questions about the errors. Teachers may share and discuss with students the portfolio scoring criteria and samples of student work and may discuss a student's best pieces and choices for inclusion in the portfolio. Teachers may assist students in identifying a variety of tasks that address the required types of portfolio entries and may assign peer tutors and others to assist students with portfolio development. Students must write, type or word process portfolio pieces by themselves, unless otherwise allowed as accommodations.

Teachers and others may not provide any assistance that diminishes personal ownership of the portfolio and may not alter documentation attesting that the portfolio contents were produced by the student. No one shall make direct corrections or revisions of portfolio entries except for the student. No changes shall be made to portfolio contents after the completion date.

Writing Portfolio Scoring. Only certified school personnel who have received current KDE training may provide accountability scores. Scorers should use current scoring materials and apply the scoring standards accurately and consistently. Scoring judgments are made on the basis of the scoring guide, benchmarks and reference to high-end portfolios to resolve decisions about performance level. The district shall maintain documentation that all scorers of writing portfolios have been appropriately trained. No individual shall instruct or encourage teachers to assign higher or lower scores than warranted and scoring accuracy should not be compromised by lack of adequate training or inappropriate scoring conditions.

Inclusion of Special Populations. An individual who provides any accommodation to a student with disabilities on any component of the statewide assessment shall be trained in his role and responsibilities and abide by confidentiality laws, the Administration Code, and the conditions as described in the student's IEP, 504 Plan, or LEP Plan. Any accommodations or modifications shall also be consistent with 703 KAR 5:070.

Alternate Portfolios. A student who meets all the eligibility requirements for the Alternate Portfolio Program may submit an alternate portfolio. Any intervention from teachers, peers or others should enhance a student's ownership of her portfolio. Teachers, parents, friends, and peers may assume support roles as listeners, responders, and encouragers. Only certified school personnel who have received current KDE training may provide accountability scores. Scorers should refer to the terms used in the Alternate Portfolio Program Holistic Scoring Guide and score only evidence seen. The district shall maintain documentation showing that scorers have been appropriately trained. No teacher-authored materials shall be included other than the entry cover page. The student may use an accommodation or assistive device only if it is a regular part of that student's instruction. No additions, subtractions, or revisions may be made after the completion deadline.

Violations of the
Administrative Code

III. Violations of the Administration Code for Kentucky's Educational Assessment Program.

The following steps shall be taken for any alleged state testing violation:

1. An allegation of inappropriate testing practices received at KDE shall be referred to the Bureau of Management Support Services, Division of Management Assistance (DMA).

2. DMA staff shall manage the process for investigating each allegation of inappropriate testing practice.
3. DMA staff shall report all findings for each allegation to the Board of Review consisting of members appointed by the Commissioner of Education (COE).
4. The Board of Review shall review the findings and make a recommendation to the COE.
5. The COE shall make a final determination and then notify the school district superintendent and the school board chairperson. If an allegation is determined to be valid and warrants invalidation or change of scores, the COE shall direct the Deputy Commissioner of Learning Support Services to make appropriate adjustments in a school or district's scores. If it appears a school district employee is guilty of wrongdoing, within 45 days, the local district superintendent shall report in writing to the COE whether disciplinary action was taken or considered necessary and shall comply with his reporting responsibility pursuant to KRS 161.120. If school or district accountability indices are adjusted as a result of the Commissioner's final determination, individual student reports shall not be changed, but changes to school or district accountability indices shall be reflected. Scores used to calculate the affected growth indices shall be adjusted, and may be reduced to non-performance for accountability purposes.
6. After the local district receives the letter from the Commissioner of the action to be taken by the Department, the school may challenge the action by appealing the next performance judgment it receives, as described in 703 KAR 5:050.

IV. Review of Secure Assessment Components by Local District and Other Certified Staff, Parents, and Persons Not in the Employment of a Kentucky Public School District.

While KDE does not require individual student participation in the statewide testing program, KDE shall hold schools accountable for the performance of all students. In the absence of assessment information about the performance of a student, the school shall be assigned a non-performance (low novice) level for that student.

Local district and other certified staff shall not be permitted routine and systematic access to the assessment. If a district chooses to assist in the review of secure testing materials, the review shall take place in the presence of the local district assessment coordinator. If a district chooses not to permit the review of secure materials under its auspices, KDE may permit review based on the availability of appropriate staff to supervise the review activities.

Reporting of Nonacademic Indicators

V. Proper Reporting of Nonacademic Indicators (Attendance, Retention, Dropout, and Transition to Adult Life).

Local districts shall be responsible for submitting this data as accurately as possible and are responsible for informing KDE of any known errors in the data reported. Reporting incorrect data shall be considered a violation of the Administration Code and shall be treated as described in Section III.

Assistance to Schools, Scholastic Audits

703 KAR 5:120 establishes standards for assistance to schools and for conducting scholastic audits.

Section 1. The following terms are defined:

1. Assistance line;
2. In need of assistance;
3. Level 1;
4. Level 2;
5. Level 3;
6. Progressing;
7. Sample of schools;
8. Scholastic audit;
9. School classification;
10. School improvement plan;
11. School portfolio;
12. Standard error of measurement; and
13. Standards and indicators for school improvement.

Section 2. A level 1 school shall conduct a scholastic review and self-study facilitated by the district's professional development coordinator with assistance provided by KDE staff. A Level 1 school may be eligible to receive Commonwealth school improvement funds.

Section 3. A level 2 school shall receive a scholastic review facilitated and chaired by a designee of the Commissioner of Education with assistance from the district's central office staff. A Level 2 school may be eligible to receive Commonwealth school improvement funds.

Section 4. A Level 3 school shall receive education assistance from a highly skilled educator under KRS 158.782 and a scholastic audit. A Level 3 school shall be eligible to receive Commonwealth school improvement funds.

Section 5. Evaluation of school personnel in a Level 3 school shall address specific issues, including the district's evaluation plan and process for certified staff, and the need for additional staff evaluations.

Section 6. If a school is classified as a Level 3 school for two consecutive biennia, a student attending the Level 3 school may transfer to a school with an accountability index above its assistance line. The superintendent shall select the receiving successful school in the home district or make arrangements with a neighboring district. The school district in which the student is enrolled shall retain the SEEK funding and the student's resident district shall be responsible for all transportation costs incurred as a result of a student transferring.

Section 7. If a school is classified as Level 3, a scholastic audit team may request the COE to recommend to a local board of education the removal of a school council member under KRS 160.347.

Section 8. Members of the scholastic audit team shall be selected and trained from a pool of candidates who have submitted an application to the KDE. This section of the regulation describes topics that must be included in the training the team members shall receive, including developing a comprehensive school improvement plan, building capacity for school leadership, organizing the school to maximize use of resources, developing an effective learning community, conducting professional growth and evaluation of certified personnel, and assessing and advising compliance with statutes and regulations.

The scholastic audit team shall consist of the following six members from which the COE shall name the chairperson:

1. A highly skilled educator;

2. An active or retired teacher from another district;
3. An active or retired principal from another district;
4. An active or retired administrator from another district;
5. A parent or legal guardian; and
6. An active or retired university faculty member.

Prior to the scholastic audit, the school principal shall prepare a school portfolio for use in creating a profile of the strengths and limitations of the school's instructional and organizational effectiveness. The scholastic audit team shall evaluate a school's learning environment, efficiency, and student academic performance by using "Standards and Indicators for School Improvement". The audit team shall make recommendations for assistance, share a draft report with the school faculty and council members prior to departure, and submit a final exit report within three weeks following the site visit.

The school principal and other school council members shall notify parents and interested community members of the findings and recommendations of the audit team. The audit findings shall be presented and discussed on the agenda of the next school council meeting and at a local board of education meeting.

School improvement plans shall be based on recommendations from the audit team's exit report and research-based standards and indicators of quality. Amending a school plan shall be a local decision of which the district is notified.

Section 9. A principal of a school classified as Level 1, 2, or 3 shall participate in at least twelve hours of professional development activities within twelve months of the classification of the school.

Section 10. The KDE shall conduct scholastic audits in a random sample of schools.

Section 11. The "Standards and Indicators for School Improvement" is incorporated by reference. It establishes nine standards with multiple indicators relevant to the attainment of each standard. The nine standards are:

1. Academic Performance – Curriculum
2. Academic Performance – Classroom Evaluation/Assessment
3. Academic Performance – Instruction
4. Learning Environment – School Culture
5. Learning Environment – Student, Family and Community Support
6. Learning Environment – Professional Growth, Development & Evaluation
7. Efficiency – Leadership
8. Efficiency – Organizational Structure and Resources
9. Efficiency – Comprehensive and Effective Planning

District Rewards

703 KAR 5:130 establishes eligibility for district rewards, and it establishes procedures for determining assistance and consequences for local school districts having schools in need of assistance as defined in 703 KAR 5:020.

Dropout data generated at an A2 - A6 school shall be attributed to the school district in which the A2 – A6 school is located unless the district can identify the A1 school that the student would have attended. In that case, the dropout data shall be assigned to the A1 school.

A local district in which all schools are classified as progressing or meets goal under 703 KAR 5:020 and meets the dropout criteria to earn rewards in 703

KAR 5:020 shall be declared an exemplary growth district and shall receive rewards determined by the KBE.

A local school district shall be held accountable for providing its schools appropriate instructional leadership and instructional support. A local school district containing a school classified as Level 3 shall modify its district consolidated plan by including a specific support plan designed to assist each Level 3 school in improving its academic achievement. If a school is classified as Level 3 for 2 or more consecutive accountability cycles, the school district shall be subject to a district audit conducted by a district evaluation team.

A local school district shall address particular areas in its school support plan, including instructional staff access to curriculum-related materials and training, professional development planning process, structure for instructional improvement, financial services and support, adequate maintenance of facilities, and an effective certified evaluation program.

The district evaluation team shall submit a report, including its recommendations, to the COE, the district superintendent, and the local board of education within two weeks of its review. The report shall be presented by a member of the district evaluation team at a local board of education meeting with opportunity for public comment.

School and District Report Cards

703 KAR 5:140 establishes the standards for a school and district report card.

The following terms are defined:

1. Average student/teacher ratio;
2. Average years of experience;
3. Base year;
4. Certified teacher;
5. Content-focused professional development;
6. District report card (base);
7. District report card (expanded);
8. School;
9. School report card (base);
10. School report card (expanded);
11. School safety data;
12. Spending per student – district;
13. Spending per student – school;
14. Spending per student – state; and
15. Total enrollment.

A school report card (base) shall be sent to the parents or guardians of each student in a school by U.S. mail, unless a waiver is granted by the KDE allowing a school to use a method of distribution that is equally effective. A school report card shall provide specific required information including the following: relevant contact information, the school's total enrollment, results of all components of CATS, teacher qualification information, school safety data, student resource data, parental involvement information, and a narrative describing actions being taken to address issues regarding the equity of the delivery of educational services to all students.

A school report card (expanded) shall be available for viewing on request in the office of the school. It shall include the specific information required by this regulation, including the following: data disaggregation pages, documentation of plans for assisting students at risk of failure, number of students participating

in special education programs, number of students and percentage of student population receiving accommodations, executive summary from the school consolidated plan, a listing of average class size, a school technology report, and the number of students enrolled for a fifth year in the primary program.

Upon the implementation of a statewide student database, the expanded school report card shall include additional information regarding the advanced placement subjects offered by a high school, an indication of whether a Commonwealth diploma is offered, and the total number of students enrolled in the gifted and talented program.

A district report card (base) shall include a district level summary of all school data required on the school report card (base) and shall be the aggregation of the school report cards by grade level.

A district report card (expanded) shall be available for viewing in the district central office. It shall include information required by the regulation, including the following: data disaggregation pages, documentation of plans to assist students at risk of academic failure, number of students participating in special education, number of students receiving instructional accommodations, executive summary from the district consolidated plans, average class sizes, technology report, number of students enrolled for a fifth year in the primary program, and copies of all base school report cards.

Upon implementation of a statewide student database, the expanded district report card shall include additional information as described above in the expanded school report card.

The school council shall review and approve the school report card (base and expanded) before it is printed. A school report card (base) shall be printed and sent to all parents no later than 77 calendar days from the release of data to the schools. A school report card (expanded) shall be available in the schools no later than the same date.

KDE shall make district and school data available electronically no later than November 1 of each year. A district has 21 days to report inaccuracies and request the data be changed. The KDE then has 21 days to correct the data or determine that no change will be made.

A district report card (base) shall be published in the newspaper with the largest circulation in the county no later than the second Sunday in February. A district report card (expanded) shall be available in the district central office no later than that same date.

KDE shall conduct an audit of school and district report cards for compliance with this regulation. If a school district fails to meet the timelines for publication, it shall communicate by letter to the KDE and identify the component which has not been published and indicate when it was or will be communicated to the appropriate public. If a school district intentionally publishes incorrect information, alters data, or refuses to produce a required component of a school report card, the matter shall be referred to the Division of Management Assistance.

The "Calculation Procedures for Data Included in the School Report Card", August, 2000 is incorporated by reference. This document is a compilation of all calculation procedures used in the school report card components.

Appendix B

Kentucky Core Content Test Sample Questions

Grade 4 Reading

- Multiple Choice: Which words BEST describe Matthew at the beginning of the story? (based on preceding reading passage)
- A. upset and impatient
 - B. cheerful but bored
 - C. anxious and frightened
 - D. puzzled but curious
- Open Response: In the story, “First Light,” Matthew woke up in another time period, the 1850s.
- A. Describe FOUR things Matthew discovered that were different from what he was used to in his present life.
 - B. Explain how each of those differences affected him. Use information from the story to explain your answer.

Grade 4 Science

- Multiple Choice: When water changes from solid to liquid to gas, which statement is true?
- A. The mass (amount) stays the same.
 - B. The temperature goes down.
 - C. The temperature stays the same
 - D. The mass (amount) is greater.
- Open Response: In spring, ice and snow change and become liquid water. In summer, when it rains on the sidewalk, the water “dries up” and becomes water vapor, which is gas.
- A. Tell what happens to make the ice and snow become liquid and the water on the sidewalk become vapor.
 - B. Give TWO examples of ways that these changes to water are important in people’s lives.

Grade 5
Social Studies

- Multiple Choice: A means of communication that is used in some homes today but was not available 20 years ago is the
- A. television
 - B. telephone
 - C. stereo
 - D. computer
- Open Response: Pioneers in the 1700s and 1800s did not go to movies, watch television, or play video games.
- A. Describe TWO forms of entertainment that were popular during “pioneer times” but are not as popular today.
 - B. Explain why they are not as popular today.

Grade 5
Mathematics

- Multiple Choice: Brittany set a school record by jumping rope 3,618 times. Andrew is trying to tie her record. So far, he has jumped 1,909 times. How many more times must he jump to tie her record?
- A. 1,709
 - B. 1,711
 - C. 2,309
 - D. 2,311
- Open Response: Corina was investigating information about natural wonders of the world.
- She found that Mt. Everest is the highest mountain in the world. It is 29,028 feet ABOVE sea level.
 - She found that the Marianas Trench in the Pacific Ocean is the lowest point on Earth. It is 35,840 feet BELOW sea level.
- A. If Corina could throw a rock from the top of Mt. Everest to the bottom of the Marianas Trench, how many feet would it fall?
 - B. Draw a diagram and explain your answer for **Part A**.

Grade 7
Reading

- Multiple Choice: According to the article, one strategy you should use before you begin to read nonfiction is to (based on preceding reading passage)
- A. separate facts from opinion
 - B. evaluate the material
 - C. question and build an understanding
 - D. preview the selection
- Open Response: In this excerpt, the author describes Dolley Madison's actions before the British burned Washington D.C. (based on preceding reading passage)
- A. Identify **two** character traits of Mrs. Madison.
 - B. Describe the actions that illustrate her character traits. Use details from the article to support your answers.

Grade 7
Science

- Multiple Choice: In which part of a plant does photosynthesis take place?
- A. bark
 - B. flowers
 - C. leaves
 - D. roots
- Open Response: Plants and animals rely on one another for the production of oxygen and carbon dioxide.
- A. Describe this relationship
 - B. Use a diagram or flow chart to illustrate your description of the relationship.

Grade 8
Mathematics

- Multiple Choice: Evaluate the expression:
 $3a - 2(b^9)$ where $a = 5$ and $b = 6$.
- A. -15
 - B. -10
 - C. 15
 - D. 30

- Open Response: Martin said, “I am thinking of a whole number between 100 and 300.
- The number is divisible by 3 but not by 9.
 - The ones digit is the sum of the hundreds digit and the tens digit.”
- A. Show why 153 cannot be Martin’s number.
- B. Find all the numbers that match Martin’s clues. Show all your work.
- C. ~~Me~~ one more clue that would limit the answer in **Part B** to one and only one correct number.

Grade 8
Social Studies

- Multiple Choice: Apartheid was the South African policy of separating people based on their race. Policies such as these promote
- A. peace
- B. discrimination
- C. social interaction
- D. equality
- Open Response: Social, economic, and cultural differences developed between the North and South in the early 1800s.
- A. Describe **two** differences between the North and the South in this period.
- B. Explain how these differences led to the Civil ~~W~~.

Grade 10
Reading

- Multiple Choice: The expression “the same thing she had said at least a thousand times”. is an example of (based on preceding reading passage)
- A. literal language
- B. symbolism
- C. hyperbole
- D. persuasion
- Open Response: An author organizes information in a particular way to convey important ideas. (based on preceding reading passage)
- A. ~~Not~~ are the three main recommendations in this article?
- B. Explain how these three recommendations are organized and presented.

**Grade 11
Mathematics**

- Multiple Choice: You have defined a new operation such that $a \star b = ab$. Which expression is equal to $b \star b$?
- A. ab^2
 - B. bab
 - C. $2b^2$
 - D. b^2
- Open Response: Jamie, Chris, and Pat are outlining a square foundation for a storage building. They have a string, tape measure, and a protractor. Each person's method for forming a square is given below:
- Jamie's method: "Cut four strings that have the same length as the sides of the square storage building. Place these strings to form a quadrilateral. That quadrilateral will be a square."
- A. ~~W~~ Jamie's method **always** form a square. Justify your reasoning using the properties of squares.
- Chris' method: "Cut four strings that have the same length as the sides of the square storage building. Place these strings to form a quadrilateral, making sure that two of the adjacent sides form a right angle."
- B. ~~W~~ Chris' method **always** form a square. Justify your reasoning using the properties of squares.
- Pat's method: "Cut two strings the same length as diagonals of the square base of the storage building. Fold them in half, marking the center of each string. Unfold the strings and place them on the ground so that they intersect at their centers to form an X. Connect the endpoints of the string to form a quadrilateral. That quadrilateral will be a square."
- C. ~~W~~ Pat's method **always** form a square. Justify your reasoning using the properties of squares.

**Grade 11
Science**

- Multiple Choice: Which statement about DNA is correct?
- A. A child's DNA will be unrelated to the mother's or father's DNA.
 - B. A child's DNA will show similarities to both the mother's and father's DNA.
 - C. A female child's DNA will exactly match the mother's DNA.
 - D. A male child's DNA will exactly match the father's DNA.

Open Response: The diagram above shows a cell with its organelles. Select four organelles from the diagram and explain how the structures and functions of those organelles within the cell are similar to the structures and functions of the different parts of your school.

Grade 11
Social Studies

Multiple Choice: During the Renaissance, there was a rebirth of interest in, and the study of, the classical cultures of

- A. Greece and Rome
- B. Japan and India
- C. Egypt and Turkey
- D. Russia and China

Open Response: For many years, countries have released industrial and human waste into the ocean. In recent years, scientists have begun to see changes in the ocean's ecosystem due to contamination by this waste.

- A. Identify **three** ways that changes in the ocean's ecosystem might affect human life.
- B. Discuss **two** possible strategies to help solve the problems associated with the contamination of the ocean's ecosystem.

Appendix C

“Needs Assistance” Schools

Assistance Provided to Level 1, Level 2, and Level 3 Schools

As discussed in Chapter 2 of this report, a school that has an accountability index that falls below its assistance line is classified as “needs assistance.” The Kentucky Department of Education divides the schools that need assistance into thirds and ranks them as Level 1, Level 2, and Level 3. The assistance provided differs by level.

Level 1 Assistance

Schools in the Level 1 assistance classification are required to undergo a scholastic self-review by a team set up by the local school district. In accordance with Section 2 of 703 KAR 5:120, Level 1 schools must conduct a scholastic review and self-study facilitated by the district’s professional development coordinator, with assistance provided by the Kentucky Department of Education (KDE) staff.

The Commissioner of Education appoints the chair of the scholastic self-review team in consultation with the district superintendent. The chair oversees the process and drafts the postaudit report.

The scholastic self-review team evaluates the school’s academic performance, learning environment, and efficiency using the “Standards and Indicators for School Improvement.” The nine standards on which the scholastic review and audit team evaluates a school are listed below. Each standard consists of several measurable indicators.

Academic Performance

1. Curriculum: rigorous, intentional, and aligned to state and local standards.
2. Classroom Evaluation and Assessment: multiple evaluation and assessment strategies.
3. Instruction: instructional program engages all students.

Learning Environment

4. School Culture: effective learning community, supports a climate conducive to performance excellence.
5. Student, Family, and Community Support: school works with families and community to remove barriers to learning.
6. Professional Growth, Development, and Evaluation: research-based professional development and performance evaluation to improve teaching and learning.

Efficiency

7. Leadership: instructional decisions focus on support for teaching and learning, organizational direction, high performance expectations, creating a learning culture, and developing leadership capacity.
8. Organizational Structure and Resources: organized to maximize use of all available resources to support high student and staff performance.
9. Comprehensive and Effective Planning: school improvement plans.

Once the self-review is complete, the team makes recommendations to improve teaching and learning. These recommendations are for inclusion in the school's existing comprehensive improvement plan; however, Level 1 schools are not required to adopt the recommendations.

Level 2 Assistance

Schools that fall into the Level 2 assistance classification are required to undergo a scholastic review by a team set up by KDE. In accordance with Section 3 of 703 KAR 5:120, a designee of the Commissioner of Education facilitates the scholastic review of Level 2 schools with assistance from the district's central office staff.

A Level 2 scholastic review by KDE consists of the following:

- Use of the "Standards and Indicators for School Improvement" to evaluate the school's academic performance, learning environment, and efficiency;
- Formulation of recommendations to improve teaching and learning for inclusion in the existing consolidated school improvement plan; and
- Review of the district's certified employee evaluation plan with development of recommendations for implementing a professional growth and evaluation plan.

After the review, the principal of a Level 2 assistance school must inform parents and community members of the findings and recommendations of the scholastic review team. The findings are presented and discussed at the school council and local board of education meetings. The school is not required to implement the recommendations.

Level 3 Assistance

Schools that fall into the Level 3 assistance classification are required to undergo a scholastic audit by an external team coordinated by KDE. The audit process is similar to the review of Level 2 schools; however, Level 3 schools are required to implement the recommendations made by the external auditors. Level 3 schools must also receive education assistance from a highly skilled educator.

Highly Skilled Educators. As a component of the Kentucky Education Reform Act, Kentucky began identifying and training experienced educators to work with schools in need of assistance. In 1998, the General Assembly revisited this feature of education reform and established the Highly Skilled Educators (HSEs) program that is currently utilized under CATS.

Under the HSE program, the Commissioner of Education and an education advisory committee use an open application process to identify experienced teachers and administrators with strong organizational, curricular, and interpersonal skills. Each year around 60 HSEs are selected and assigned to low-performing schools across the state. HSEs receive training from the Kentucky Department of Education prior to assessing low-performing schools. These educators may receive up to two years of paid leave from their local districts and are paid at 135 percent of their local salaries while they are performing as a highly skilled educator.

HSEs use the same “Standards and Indicators for School Improvement” that are used in the Level 1 and Level 2 reviews to evaluate Level 3 schools’ academic performance, learning environment, and efficiency. Under KRS 158.782, the HSE program is designed to support improved teaching and learning by establishing the following:

- Criteria for identifying successful strategies of assistance;
- Policies and procedures for providing education assistance, which may include training, making assignments, employing certified personnel, and setting salaries that may include supplements; and
- Duties of those providing education assistance, which may include personnel evaluation and recommendations concerning retention, dismissal, or transfer of personnel.

The HSE program will receive \$5.8 million in funding in fiscal year 2003 and \$5.7 million in FY 2004. The Commissioner of Education has discretion in the use of HSE funds. Specifically, Kentucky's 2002-04 state budget allows HSE funds to be used for other "intervention services that may be required" by the federal No Child Left Behind Act.

Level 3 schools are required to implement the recommendations of the external auditors. After the external auditors make recommendations, the highly skilled educators assist schools in complying with the audit recommendations.

School Improvement Funds

Under KRS 158.805, all schools classified as needs assistance are eligible to receive school improvement funds. The Commonwealth School Improvement Fund was created to assist local schools in pursuing new and innovative strategies to meet the education needs of students and to raise school performance levels.

The priority for use of school improvement funds for school years 2002-2003 and 2003-2004 is to provide technical assistance to reduce the achievement gaps among various subgroups of students. The improvement funds are also to be used to address the recommendations of the scholastic review or audit.

In accordance with KRS 158.805, the Kentucky Board of Education has sole authority to develop criteria for and approve rewarding of school improvement grants or awards. The amount each school receives is based on a base amount calculated using average daily attendance. In addition to the base amount calculated for each school, the remaining funds are allocated based on identified criteria that reveal the school's degree of need. For the 2002-2004 biennium, the additional allocation is based on the size of school achievement gaps—the greater the achievement gaps, the more improvement money the school receives.

Pursuant to KRS 158.805 school improvement money comes directly out of the General Fund. A total of \$2,054,200 was expended out of the General Fund in FY 2003 for school improvement.

Schools receive improvement money upon approval of their master agreement. Master agreements are approved annually for districts with schools that are in need of assistance. District master agreements are not processed for approval until budgetary information is received from schools and districts. Funds are distributed as soon as the Finance and Administration Cabinet and the Legislative Research Commission approve the master agreements. Improvement funds are required to be expended by June 30 of the year of receipt, with the exception of 10 percent of a school's improvement dollars, which the school is allowed to carry over to the next year.

Notifications of school improvement fund awards were sent to schools in December 2002. Eligible school districts' actual receipt of their school improvement money varied due to the master agreement process. For the 2002-2003 school year, districts received their improvement funds between March 24 and June 17, with the majority receiving their funds in late May or early June.

Once a needs assistance school receives improvement money, those funds may be used on any activity addressed in the school's comprehensive school improvement plan. Improvement funds can be used as supplemental pay for teachers who take part in professional development opportunities, but schools are restricted from using the money to pay teacher salaries. The school may spend up to 20 percent of the improvement funds to purchase equipment or technology as long as the district matches the expenditure. No other restrictions are placed on the improvement funds received by schools.

Schools that receive improvement funds are not formally audited to ensure funds are spent in accordance with program guidelines, but the technical help provided to needs assistance schools is intended to provide some measure of spending oversight. The assistance is narrowly tailored to help schools reduce achievement gaps and implement recommendations made during the review and audit process.

Appendix D

The Surveys of Teachers, Principals, and Superintendents

This appendix details how the surveys were developed and conducted and provides evidence that the samples may be considered representative of teachers, principals, and superintendents in Kentucky's public school system. Frequency tables for teachers', principals', and superintendents' answers to the questions are also included.

How the Questionnaires Were Developed

The Kentucky Department of Education (KDE) provided lists of superintendents, principals, and teachers for the 2002-2003 school year. Program Review staff developed separate questionnaires for each of the three groups. To avoid bias in the wording of the questions, two versions of the teachers' and principals' questionnaires were distributed: a positively worded version and a negatively worded version. For example, both teachers and principals were asked:

What is your opinion of the Core Content for Assessment?

Half the teachers and principals received the first group of responses and half received the second:

Positive Version

Very Positively
Somewhat Positively
Neutral or No Effect
Somewhat Negatively
Very Negatively

Negative Version

Very Negatively
Somewhat Negatively
Neutral or No Effect
Somewhat Positively
Very Positively

There were no significant differences in responses based on the version of the questionnaire received by teachers and principals. Only the negatively worded version was used for the survey of superintendents, but in the frequency tables in this appendix, the response categories are listed in the same order as for the teacher and principal survey tables for ease of comparison.

Potential survey questions were developed based on:

- A review of CATS assessment materials;
- Interviews with KDE staff and school administrators; and
- Teacher focus groups.

Staff used the information gathered to identify key issues and concerns regarding the accountability testing system. Staff sent a draft questionnaire to teachers who participated in the focus groups. Their comments were incorporated into the questionnaires.

How the Surveys Were Conducted

Teachers. A list of Kentucky's 40,360 public school teachers was sorted by type of school, and a random sample was taken from each of the school categories. A four-page questionnaire was then mailed to 2,100 Kentucky public school teachers selected for the sample. The number of teachers taken from each category was representative of the proportion that the category represented for teachers as a whole.

School Type	Number of Teachers	Percent of Total	Sample
Elementary	19,672	48.7%	1,024
Middle/Jr. High	8,152	20.2%	424
High School	11,166	27.7%	581
Middle and High	94	0.2%	5
Alternative	627	1.6%	33
Voc/Technical	192	0.5%	10
Primary-12	157	0.4%	8
K-8	300	0.7%	16
Total	40,360	100.0%	2,100

About 10 days after the questionnaire was mailed, each teacher in the sample was mailed a postcard asking him or her to respond to the survey or thanking him or her for doing so already. Those who did not respond to the first mailing or postcard were then mailed a second copy of the questionnaire. Seven hundred and ninety-nine teachers returned surveys for a response rate of 38 percent. Approximately half the surveys returned were the positively worded version and half were the negatively worded version.

Principals. A web-based survey was conducted to capture the opinions of school principals regarding the CATS assessment. An email distribution list was compiled of Kentucky's 1,238 principals. All principals were sent an email asking them to complete the online questionnaire and giving instructions.

Principals were also sent a follow-up email one week after the first email requests were sent.

Of the 1,238 principals, 515 responded to the web-based survey for a response rate of 42 percent. The completed questionnaires were evenly divided between the two versions.

Superintendents. A web-based survey was conducted to capture the opinions of school superintendents regarding the CATS assessment. An email distribution list was compiled of each district's superintendent for Kentucky's 176 school districts. All superintendents were sent an email asking them to complete the entire questionnaire and giving them instructions.

Superintendents were sent a follow-up email one week after the first email request, either thanking them for responding already or again asking them to do so.

Of the 176 superintendents, 105 responded to the web-based survey for a response rate of 72 percent.

The Representativeness of the Sample

It cannot be ruled out that the educators who chose to respond to the surveys hold meaningfully different opinions from those who did not. It is possible, however, to analyze available information to increase confidence that this sample is representative of teachers, principals, and superintendents in Kentucky.

One way to address the question of response bias is to compare certain characteristics of the sample to those of the total population. As shown in the tables below, the sample of teachers seems to be a reasonable representation of all public school teachers.

School Type for Survey Respondents and All Teachers

School Type	Number of Schools	Percent of Total	Sample	Sample Percent of Total	Number of Respondents	Percent of Respondents
Elementary	19,672	48.7%	1024	48.8%	393	49.2%
Middle/Jr. High	8,152	20.2%	424	20.2%	180	22.5%
High School	11,166	27.7%	581	27.7%	198	24.8%
Middle and High	94	0.2%	5	0.2%	2	0.3%
Alternative	627	1.6%	33	1.6%	13	1.6%
Voc/Technical	192	0.5%	10	0.5%	3	0.4%
Primary-12	157	0.4%	8	0.4%	4	0.5%
K-8	300	0.7%	16	0.8%	6	0.8%
Total	40,360	100.0%	2,100	100%	799	100.0%

Educational Region for Survey Respondents and All Teachers

Region	Total	Percent of Total	Sample	Sample Percent of Total	Number of Respondents	Percent of Respondents
1	6,337	15.7%	248	11.9%	117	14.7%
2	7,628	18.9%	325	15.5%	150	18.8%
3	5,328	13.2%	271	14.2%	103	12.9%
4	5,247	13.0%	290	13.8%	107	13.4%
5	5,328	13.2%	357	16.3%	114	14.3%
6	5,004	12.4%	294	13.8%	101	12.7%
7	2,704	6.7%	167	8.1%	63	7.9%
8	2,784	6.9%	148	6.4%	44	5.5%
Total	40,360	100%	2,100	100.0%	799	100%

Another way to address the question of potential bias is to compare those who responded to the survey quickly to those who responded later. The logic is that if there is a response bias, those who responded later may be similar to those who did not respond at all. For example, a worry with most surveys is that those who have especially strong attitudes about the topic of the survey are more likely to respond with little prompting. If early respondents are very different from late respondents, that could indicate response bias. To see if that is the case here, the surveyed educators are divided into those who responded after receiving the questionnaire in the mail once, and those who did not respond until after the final reminder was mailed. The answers of late respondents are then used as proxies for those who did not respond to the survey at all. Based on this assumption, it is possible to project what the survey results would be if the response rate was 100 percent.

The tables below compare the actual results from the survey with the projected results using the answers of late responders as proxies for nonresponders. The actual and projected results are shown for three questions regarding the CATS assessment asked of teachers, principals, and superintendents. The differences between the actual survey of results and the projected results are small. The representativeness of a sample cannot be guaranteed, but all indicators suggest that those who responded to the survey are typical educators in terms of their views on the CATS assessment.

Teachers

What is your opinion of the Core Content for Assessment?

	Survey	Projected
Very Satisfied	10%	10%
Somewhat Satisfied	42%	41%
Neutral	21%	23%
Somewhat Dissatisfied	20%	20%
Very Dissatisfied	7%	6%

Do you agree or disagree that the consequences to schools that fail to improve are appropriate?

	Survey	Projected
Strongly Agree	2%	2%
Agree	18%	20%
Disagree	38%	36%
Strongly Disagree	20%	18%
Not Sure	22%	24%

Is the weight of writing portfolios on the accountability index appropriate for the amount of instructional time each requires at your school?

	Survey	Projected
About Right?	36%	36%
Too High?	48%	49%
Too Low?	16%	15%

Principals

What is your opinion of the Core Content for Assessment?

	Survey	Projected
Very Satisfied	6%	5%
Somewhat Satisfied	15%	13%
Neutral	6%	6%
Somewhat Dissatisfied	55%	58%
Very Dissatisfied	18%	18%

Do you agree or disagree that the consequences to schools that fail to improve are appropriate?

	Survey	Projected
Strongly Agree	3%	3%
Agree	30%	32%
Disagree	39%	37%
Strongly Disagree	18%	18%
Not Sure	10%	10%

Is the weight of writing portfolios on the accountability index appropriate for the amount of instructional time each requires at your school?

	Survey	Projected
About Right?	42%	43%
Too High?	31%	32%
Too Low?	26%	25%

Superintendents

What is your opinion of the Core Content for Assessment?

	Survey	Projected
Very Satisfied	15%	16%
Somewhat Satisfied	69%	70%
Neutral	4%	4%
Somewhat Dissatisfied	11%	8%
Very Dissatisfied	1%	1%

Do you disagree or agree that the consequences to schools that fail to improve are appropriate?

	Survey	Projected
Strongly Agree	11%	10%
Agree	44%	54%
Disagree	32%	25%
Strongly Disagree	7%	4%
Not Sure	7%	7%

Is the weight of writing portfolios on the accountability index appropriate for the amount of instructional time each requires at your school?

	Survey	Projected
About Right?	49%	57%
Too High?	38%	31%
Too Low?	13%	12%

**Survey of Teachers
Responses to Questions**
(Frequencies Provided for Closed-ended Questions)

1: What grade(s) do you teach?

2: How many years have you been teaching at your current school?

3: What subject(s) do you teach?

4: Please select all that apply to you.

I have never participated in scoring writing portfolios.	212	27%
I do not participate in scoring writing portfolios at this time but have in the past.	188	24%
Writing Portfolio Scorer	391	51%
Regional Writing Consultant	9	1%
Regional Writing Coordinator	8	1%
Writing Portfolio Cluster Leader	50	6%
Scoring Accuracy Assurance Team Member	24	3%
Number of teachers answering the question	773	*

*More than one response could be checked, total percent adds to over 100%

5: What is your opinion of the Core Content for Assessment?

Very Satisfied	79	10%
Somewhat Satisfied	322	42%
Neutral	162	21%
Somewhat Dissatisfied	157	20%
Very Dissatisfied	52	7%
Total	772	100%

Questions 6-10 relate to testing and the core content.

6: Overall, how does CATS affect what students learn?

Very Positively	62	8%
Somewhat Positively	362	47%
Neutral or No Effect	138	18%
Somewhat Negatively	174	23%
Very Negatively	34	4%
Total	770	100%

7: Do some types of students benefit more from CATS than others?

Yes	457	59%
No	151	19%
Not Sure	171	22%
Total	779	100%

8: Are some types of students disadvantaged by CATS?

Yes	611	79%
No	65	8%
Not Sure	101	13%
Total	777	100%

9: Do you have enough instructional time to adequately teach the core content covered by assessment?

Yes	251	33%
No	464	60%
Not Sure	53	7%
Total	768	100%

10: Compared to KIRIS, the accountability system before CATS, how has CATS affected the overall quality of education provided in your school? The quality of education provided is now:

Much Better	58	8%
Somewhat Better	172	23%
About the Same	299	39%
Somewhat Worse	56	7%
Much Worse	17	2%
Not Sure	155	20%
Total	757	100%

Questions 11-14 deal specifically with CATS testing.

11: Overall, how does the CATS test affect the way you teach?

Very Positively	67	9%
Somewhat Positively	247	32%
Neutral or No Effect	214	28%
Somewhat Negatively	201	26%
Very Negatively	36	5%
Total	765	100%

12: Most questions on the CATS test are repeated from the previous version of the test. Does this repetition of questions affect the way you prepare your students for the test?

No	371	49%
Yes, Somewhat	233	31%
Yes, A Lot	56	7%
Not Sure	98	13%
Total	758	100%

13: Do some types of schools fare better on the CATS test than others?

Yes	578	74%
No	18	2%
Not Sure	186	24%
Total	782	100%

14: Do some types of schools do worse on the CATS test than others:

Yes	539	69%
No	29	4%
Not Sure	213	27%
Total	781	100%

If your school is or has been in the “needs assistance” classification in the accountability system, please answer Question 15. If not, skip to Question 16.

15: How would you rate the assistance provided?

Very Helpful	41	18%
Somewhat Helpful	105	45%
Not Very Helpful	46	20%
Not Helpful at All	11	5%
Not Sure	28	12%
Total	231	100%

16: Have you participated in the process of developing an annual improvement plan for your school to be submitted to the Kentucky Department of Education?

Yes	466	66%
No	245	34%
Total	711	100%

**If you answered YES to Question 16, please answer Question 17.
If not, skip to Question 18.**

17: How would you rate this process?

Very Helpful	87	19%
Somewhat Helpful	260	56%
Not Very Helpful	78	17%
Not Helpful at All	28	6%
Not Sure	9	2%
Total	462	100%

Biennial improvement goals have been established for each school beginning in 2002 with intermediate targets that will eventually take each school to the goal of 100 on the accountability index by 2014.

18: Do you agree or disagree that the process used by the Kentucky Board of Education to set improvement goals was appropriate?

Strongly Agree	18	2%
Agree	137	19%
Disagree	239	32%
Strongly Disagree	139	19%
Not Sure	206	28%
Total	739	100%

19: Do you agree or disagree that the consequences to schools that fail to improve are appropriate?

Strongly Agree	16	2%
Agree	135	18%
Disagree	291	38%
Strongly Disagree	157	20%
Not Sure	168	22%
Total	767	100%

WRITING PORTFOLIOS

20: Is writing portfolio training available to all teachers at your school?

Yes	660	83%
No, only those who score portfolios	100	13%
Not Sure	31	4%
Total	791	100%

21: Is writing development training available to all teachers at your school?

Yes	644	82%
No	80	10%
Not Sure	63	8%
Total	787	100%

22: Have you received writing development training?

Yes	678	86%
No	108	14%
Not Sure	1	0%
Total	787	100%

23: Do you agree or disagree that the performance categories (novice, apprentice, proficient, and distinguished) are appropriate categories to assign to students' writing portfolios?

Strongly Agree	57	7%
Agree	477	62%
Disagree	110	14%
Strongly Disagree	65	8%
Not Sure	63	8%
Total	772	100%

24: Do teachers who score portfolios have biases that affect scores?

No	203	26%
Yes, a few teachers do	271	34%
Yes, many teachers do	109	14%
Not Sure	204	26%
Total	787	100%

25: Would you agree or disagree that the amount of time it takes teachers and students to prepare writing portfolios is appropriate to the benefit received by students?

Strongly Agree	22	3%
Agree	120	15%
Disagree	234	30%
Strongly Disagree	322	41%
Not Sure	87	11%
Total	785	100%

26: Are the weights of the following subject areas on the accountability index appropriate for the amount of instructional time each requires at your school?

<i>Writing Portfolios</i>		
About Right?	249	36%
Too High?	328	48%
Too Low?	108	16%
Total	685	100%
<i>On-demand Writing</i>		
About Right?	380	56%
Too High?	253	37%
Too Low?	42	6%
Total	675	100%
<i>CTBS-Norm Referenced Testing</i>		
About Right?	477	73%
Too High?	54	8%
Too Low?	119	18%
Total	650	100%
<i>KCCT Reading</i>		
About Right?	565	87%
Too High?	57	9%
Too Low?	28	4%
Total	650	100%
<i>KCCT Mathematics</i>		
About Right?	549	85%
Too High?	64	10%
Too Low?	30	5%
Total	643	100%
<i>KCCT Science</i>		
About Right?	526	83%
Too High?	79	12%
Too Low?	32	5%
Total	637	100%
<i>KCCT Social Studies</i>		
About right?	540	84%
Too High?	73	11%
Too Low?	28	4%
Total	641	100%
<i>KCCT Arts and Humanities</i>		
About right?	384	60%
Too High?	198	31%
Too Low?	54	8%
Total	636	100%
<i>KCCT Practical Living/Vocational Skills</i>		
About Right?	409	65%
Too High?	155	25%
Too Low?	64	10%
Total	628	100%

Nonacademic Factors

Attendance Rate

About Right?	417	66%
Too High?	164	26%
Too Low?	52	8%
Total	633	100%

Retention Rate

About Right?	361	58%
Too High?	208	33%
Too Low?	55	9%
Total	624	100%

Dropout Rate (middle and high school ONLY)

About Right?	249	57%
Too High?	164	38%
Too Low?	23	5%
Total	436	100%

Transition to Adult Life (high school ONLY)

About Right?	210	64%
Too High?	78	24%
Too Low?	41	12%
Total	329	100%

27: Should any of the above components be removed?

Yes	312	43%
No	160	22%
Not Sure	251	35%
Total	723	100%

28: Should other subject areas or nonacademic factors be included in the assessment and accountability system?

Yes	130	18%
No	350	48%
Not Sure	253	35%
Total	733	100%

29: What changes, if any, would you make to the CATS assessments or the accountability system?

30: Do you think your school can reach proficiency by 2014?

Yes	157	21%
No	358	47%
Not Sure	245	32%
Total	760	100%

If you score or have scored writing portfolios, please answer Questions 31–45.

31: Do you ever score your own students' writing portfolios for accountability purposes?

Yes	202	34%
No	378	63%
Not Sure	19	3%
Total	599	100%

**If you answered YES to Question 31, please answer Question 32.
If not, skip to Question 33.**

32: Does scoring your own students' writing portfolios affect your scoring?

Yes	55	17%
No	210	67%
Not Sure	50	16%
Total	315	100%

33: How helpful or unhelpful is the video portion of the writing portfolio scoring training provided by the Kentucky Department of Education?

Very Helpful	22	4%
Somewhat Helpful	255	44%
Not Very Helpful	169	29%
Not Helpful at All	129	22%
Total	575	100%

34: How helpful or unhelpful is the portion of the writing portfolio scoring training delivered by your school's trainer/cluster leader?

Very Helpful	263	46%
Somewhat Helpful	260	45%
Not Very Helpful	32	6%
Not Helpful at All	18	3%
Total	573	100%

35: Are there aspects of the writing portfolio scoring training that are particularly helpful?

Yes	278	49%
No	129	23%
Not Sure	157	28%
Total	564	100%

36: Does your scoring trainer utilize the Scoring Training CD-ROM during your training sessions?

Yes	192	45%
No	259	22%
Not Sure	127	33%
Total	578	100%

37: Does your scoring trainer utilize the annotated high-end portfolios during your training sessions?

Yes	495	85%
No	14	2%
Not Sure	75	13%
Total	584	100%

38: What could be done to improve the writing portfolio training?

39: How many hours of writing portfolio scoring training do you receive annually?

40: Do you agree or disagree that the amount of time it takes to score writing portfolios is appropriate for the benefit received by students?

Strongly Agree	33	6%
Agree	179	30%
Disagree	161	27%
Strongly Disagree	117	20%
Not Sure	101	17%
Total	591	100%

41: Does your school have enough teachers scoring writing portfolios?

Yes	456	77%
No	73	12%
Not Sure	65	11%
Total	594	100%

42: Which writing portfolio scoring option does your school use?

Double blind scoring	387	76%
Individual scoring with informal support/discussion as necessary	28	6%
Individual scoring/blind second scoring	41	8%
Individual scoring/selected blind second scoring	14	3%
Individual scoring/selected blind second scoring/reliability check	26	5%
Individual scoring/selected blind/informal discussion	12	2%
Total	508	100%

43: Do you agree or disagree that the scoring option your school uses yields accurate writing portfolio scores?

Strongly Agree	137	23%
Agree	311	53%
Disagree	37	6%
Strongly Disagree	14	2%
Not Sure	87	15%
Total	586	100%

44: About what percentage of the time do the first two scores given to a portfolio by scorers working independently NOT agree?

45: How often do members of your scoring team score and review Quality Control portfolios during your portfolio scoring sessions?

After every five or six portfolios scored	126	28%
At the beginning and middle of each session	160	36%
At the beginning of each session	136	31%
Never	21	5%
Total	443	100%

**Survey of Principals
Responses to Questions**
(Frequencies Provided for Closed-ended Questions)

1: What is your opinion of the Core Content for Assessment?

Very Satisfied	91	18%
Somewhat Satisfied	276	55%
Neutral	32	6%
Somewhat Dissatisfied	75	15%
Very Dissatisfied	29	6%
Total	503	100%

2: Has your school vertically and horizontally aligned its curriculum and instruction?

Yes	462	91%
No	41	8%
Not Sure	7	1%
Total	510	100%

3: Does your school periodically review the aligned program?

Yes	492	96%
No	14	3%
Not Sure	4	1%
Total	510	100%

4: Overall, how does CATS affect your school's curriculum?

Very Positively	116	23%
Somewhat Positively	270	53%
Neutral/No Effect	33	6%
Somewhat Negatively	77	15%
Very Negatively	13	3%
Total	509	100%

5: Overall, how does CATS affect what students learn at your school?

Very Positively	134	26%
Somewhat Positively	231	46%
Neutral/No Effect	39	8%
Somewhat Negatively	86	17%
Very Negatively	17	3%
Total	507	100%

6: Do some types of students benefit more from CATS than others?

Yes	326	64%
No	118	23%
Not Sure	68	13%
Total	512	100%

7: Are some types of students disadvantaged by CATS?

Yes	401	79%
No	58	11%
Not Sure	51	10%
Total	510	100%

8: Do you think that teachers have enough instructional time to adequately teach the core content covered by assessment?

Yes	179	35%
No	309	60%
Not Sure	25	5%
Total	513	100%

9: Compared to KIRIS, the accountability system before CATS, how has CATS affected the quality of education provided in your school? The quality of education provided is now:

Much Better	78	15%
Somewhat Better	188	37%
About the Same	219	43%
Somewhat Worse	18	4%
Much Worse	3	1%
Total	506	100%

10: Do some types of schools do better on the CATS tests than others?

Yes	392	77%
No	37	7%
Not Sure	82	16%
Total	511	100%

11: Do some types of schools do worse on the CATS tests than others?

Yes	360	71%
No	52	10%
Not Sure	96	19%
Total	508	100%

12: Overall, how does the CATS test affect the way teachers at your school teach?

Very Positively	86	17%
Somewhat Positively	264	52%
Neutral or No Effect	21	4%
Somewhat Negatively	124	25%
Very Negatively	11	2%
Total	506	100%

13: Most questions on the CATS test are repeated from the previous version of the test. Does this repetition of questions affect the way your teachers prepare students for the test?

No	201	39%
Yes, for a few teachers	116	23%
Yes, for many teachers	118	23%
Not Sure	78	15%
Total	513	100%

If your school is or has been in the “needs assistance” classification in the accountability system, please answer Question 14. If not, skip to Question 15.

14: How would you rate the assistance provided?

Very Helpful	53	40%
Somewhat Helpful	47	35%
Not Very Helpful	23	17%
Not Helpful at All	7	5%
Not sure	4	3%
Total	134	100%

15: Have you participated in the process of developing an annual improvement plan for your school to be submitted to the Kentucky Department of Education?

Yes	348	75%
No	119	25%
Total	467	100%

16: How would you rate the process of developing an annual plan?

Very Helpful	124	36%
Somewhat Helpful	169	49%
Not Very Helpful	45	13%
Not Helpful at All	4	1%
Not Sure	4	1%
Total	346	100%

Biennial improvement goals have been established for each school beginning in 2002 with intermediate targets that will eventually take each school to the goal of 100 on the accountability index by 2014.

17: Do you agree or disagree that the process used by the Kentucky Board of Education to set improvement goals was appropriate?

Strongly Agree	29	6%
Agree	223	44%
Disagree	142	28%
Strongly Disagree	67	13%
Not Sure	47	9%
Total	508	100%

18: Do you agree or disagree that the consequences to schools that fail to improve are appropriate?

Strongly Agree	15	3%
Agree	154	30%
Disagree	197	39%
Strongly Disagree	91	18%
Not Sure	50	10%
Total	507	100%

19: Is writing portfolio training available to all teachers at your school?

Yes	457	89%
No, only those who score portfolios	56	11%
Total	513	100%

20: Is writing portfolio training completed by all teachers at your school who score writing portfolios?

Yes	475	92%
No	39	8%
Total	514	100%

21: Do all teachers in your school receive training in writing development?

Yes	432	85%
No	76	15%
Not Sure	3	1%
Total	511	100%

22: Do teachers who score portfolios have biases that affect scores?

No	283	55%
Yes, a few teachers do	104	20%
Yes, many teachers do	33	6%
Not Sure	90	18%
Total	510	100%

23: Do teachers at your school ever score their own students' portfolios for accountability purposes?

Yes	160	32%
No	319	63%
Not Sure	28	6%
Total	507	100%

24: How helpful or unhelpful is the video portion of the portfolio scoring training provided by the Kentucky Department of Education on the scoring of writing portfolios?

Extremely Helpful	36	7%
Somewhat Helpful	295	58%
Not Very Helpful	135	27%
Not Helpful At All	41	8%
Total	507	100%

25: How helpful or unhelpful is the portion of the portfolio scoring training delivered by your school's trainer/cluster leader?

Extremely Helpful	288	57%
Somewhat Helpful	202	40%
Not Very Helpful	13	3%
Not Helpful At All	3	1%
Total	506	100%

26: What do you like best about the writing portfolio training?

27: What could be done differently to improve the writing portfolio training process?

28: Do you agree or disagree that the performance categories (novice, apprentice, proficient, and distinguished) are appropriate categories to assign to a student's writing portfolio?

Strongly Agree	61	12%
Agree	325	64%
Disagree	71	14%
Strongly Disagree	26	5%
Not Sure	25	5%
Total	508	100%

29: How many hours of portfolio scoring training does the typical teacher at your school participating in the scoring of portfolios receive annually?

30: Does your school have enough teachers scoring writing portfolios?

Yes	455	89%
No	43	8%
Not Sure	13	3%
Total	511	100%

31: Would you agree or disagree that the amount of time it takes to score portfolios is appropriate to the benefit?

Strongly Agree	43	8%
Agree	260	51%
Disagree	129	25%
Strongly Disagree	57	11%
Not Sure	23	4%
Total	512	100%

32: Do you hire substitute teachers to replace teachers who are attending portfolio scoring training?

Yes	191	38%
No	317	62%
Total	508	100%

33: Do you hire substitute teachers to replace teachers who are scoring portfolios?

Yes	276	55%
No	225	45%
Total	501	100%

34: Which portfolio grading option does your school use?

Double blind scoring	395	79%
Individual scoring/selected blind/informal discussion	10	2%
Individual scoring/selected blind second scoring/reliability check	37	7%
Individual scoring with informal support/discussion as necessary	10	2%
Individual scoring/selected blind second scoring	15	3%
Individual scoring/blind second scoring	36	7%
Total	503	100%

35: Do you agree or disagree that the scoring option that your school uses yields accurate writing portfolio scores?

Strongly Agree	178	35%
Agree	283	56%
Disagree	22	4%
Strongly Disagree	5	1%
Not Sure	17	3%
Total	505	100%

36: Are the weights of the following subject areas on the accountability index appropriate for the amount of instructional time each requires at your school?

Writing Portfolios

About Right?	212	42%
Too High?	157	31%
Too Low?	131	26%
Total	500	100%

On-demand Writing

About Right?	279	56%
Too High?	175	35%
Too Low?	47	9%
Total	501	100%

CTBS/Norm-Referenced Testing

About Right?	299	59%
Too High?	16	3%
Too Low?	190	38%
Total	505	100%

KCCT Reading

About Right?	439	87%
Too High?	37	7%
Too Low?	27	5%
Total	503	100%

KCCT Mathematics

About Right?	440	88%
Too High?	43	9%
Too Low?	19	4%
Total	502	100%

KCCT Science

About Right?	428	85%
Too High?	65	13%
Too Low?	10	2%
Total	503	100%

KCCT Social Studies

About Right?	428	86%
Too High?	60	12%
Too Low?	11	2%
Total	499	100%

KCCT Arts and Humanities

About Right?	287	57%
Too High?	178	35%
Too Low?	37	7%
Total	502	100%

KCCT Practical Living/Vocational Skills

About Right?	306	61%
Too High?	162	33%
Too Low?	30	6%
Total	498	100%

Nonacademic Factors

Attendance Rate

About Right?	358	72%
Too High?	98	20%
Too Low?	43	9%
Total	499	100%

Retention Rate

About Right?	325	66%
Too High?	143	29%
Too Low?	26	5%
Total	494	100%

Dropout Rate (middle and high school ONLY)

About Right?	325	71%
Too High?	111	24%
Too Low?	19	4%
Total	455	100%

Transition to Adulthood (high school ONLY)

About Right?	313	72%
Too High?	102	23%
Too Low?	22	5%
Total	437	100%

37: Should any of the above components be removed?

Yes	272	54%
No	130	26%
Not Sure	101	20%
Total	503	100%

38: Should other subject areas or nonacademic factors be included in the assessment and accountability system?

Yes	69	14%
No	306	62%
Not Sure	119	24%
Total	494	100%

39: What changes, if any, would you make to the CATS assessments or the accountability system?

40: Do you think your school can reach proficiency by 2014?

Yes	222	44%
No	106	21%
Not sure	174	35%
Total	502	100%

Questions 41-43 deal with the federal No Child Left Behind (NCLB) Act.

41: Do you think the schools in your district can close the following achievement gaps by 2014?

Gender Achievement Gap

Yes	346	68%
No	41	8%
Not applicable	119	24%
Total	506	100%

Disability Achievement Gap

Yes	126	25%
No	200	40%
Not applicable	180	36%
Total	506	100%

English Language Proficiency Achievement Gap

Yes	132	26%
No	95	19%
Not applicable	274	55%
Total	501	100%

Income Achievement Gap

Yes	169	33%
No	173	34%
Not applicable	163	32%
Total	505	100%

Race/Ethnicity Achievement Gap

Yes	253	50%
No	47	9%
Not applicable	206	41%
Total	506	100%

42: Do you think you can have a "highly qualified" teacher in every classroom by 2005?

Yes	298	59%
No	123	24%
Not sure	83	16%
Total	504	100%

43: Are there any provisions of the national No Child Left Behind Act that you think will be difficult for your school to achieve?

Yes	158	32%
No	129	26%
Not sure	214	43%
Total	501	100%

If YES, which provision(s) of No Child Left Behind do you think will be most difficult for your school to meet?

**Survey of Superintendents
Responses to Questions**
(Frequencies Provided for Closed-ended Questions)

1: What is your opinion of the Core Content for Assessment?

Very Satisfied	14	15%
Somewhat Satisfied	66	69%
Neutral	4	4%
Somewhat Dissatisfied	11	11%
Very Dissatisfied	1	1%
Total	96	100%

2: Overall, how does CATS affect the way teachers in your district teach?

Very Positively	17	16%
Somewhat Positively	64	62%
Neutral/No Effect	3	3%
Somewhat Negatively	18	17%
Very Negatively	2	2%
Total	104	100%

3: Overall, how does CATS affect your district's curriculum?

Very Positively	24	23%
Somewhat Positively	58	56%
Neutral/No Effect	5	5%
Somewhat Negatively	15	14%
Very Negatively	2	2%
Total	104	100%

4: Overall, how does CATS affect what students in your district learn?

Very Positively	17	17%
Somewhat Positively	63	60%
Neutral/No Effect	6	6%
Somewhat Negatively	17	16%
Very Negatively	1	1%
Total	100	100%

5: Do some types of schools do better on the CATS tests than others?

Yes	76	73%
Not Sure	21	20%
No	7	7%
Total	104	100%

6: Do some types of schools do worse on the CATS tests than others?

Yes	72	72%
No	8	8%
Not Sure	20	20%
Total	100	100%

7: Do some types of districts do better on the CATS tests than others?

Yes	69	68%
No	10	10%
Not Sure	23	23%
Total	102	100%

8: Do some types of districts do worse on the CATS tests than others?

Yes	67	66%
No	10	10%
Not Sure	25	25%
Total	102	100%

9: Most questions on the CATS test are repeated from the previous version of the test. Does this repetition of questions affect the way your teachers prepare students for the test?

No	30	29%
Yes, for few teachers	20	19%
Yes, for many teachers	39	38%
Not Sure	14	14%
Total	103	100%

10: Do you think that teachers have enough instructional time to adequately teach the core content covered by assessment?

Yes	30	29%
No	69	66%
Not Sure	5	5%
Total	104	100%

11: Compared to KIRIS, the accountability system before CATS, how has CATS affected the overall quality of education provided in your district? The quality of education provided is now:

Much Better	11	11%
Somewhat Better	41	40%
About the Same	47	46%
Somewhat Worse	4	4%
Much Worse	0	0%
Total	103	100%

Biennial improvement goals have been established for each school beginning in 2002 with intermediate targets that will eventually take each school to the goal of 100 on the accountability index by 2014.

12: Do you disagree or agree that the process used by the Kentucky Board of Education to set improvement goals was appropriate?

Strongly Agree	11	11%
Agree	54	52%
Disagree	20	19%
Strongly Disagree	8	8%
Not Sure	11	11%
Total	104	100%

13: Do you disagree or agree that the consequences to schools that fail to improve are appropriate?

Strongly Agree	11	11%
Agree	46	44%
Disagree	33	32%
Strongly Disagree	7	7%
Not Sure	7	7%
Total	104	100%

14: Do you disagree or agree that the performance categories (novice, apprentice, proficient, and distinguished) are appropriate categories to assign to a student's writing portfolio?

Strongly Agree	17	16%
Agree	61	59%
Disagree	22	21%
Strongly Disagree	0	0%
Not Sure	4	4%
Total	104	100%

15: Are the weights of the following subject areas on the accountability index appropriate for the amount of instructional time each requires in your district?

Writing Portfolios

About Right?	51	49%
Too High?	39	38%
Too Low?	14	13%
Total	104	100%

On-demand Writing

About Right?	64	62%
Too High?	26	25%
Too Low?	14	13%
Total	104	100%

CTBS/Norm-Referenced Testing

About Right?	52	50%
Too High?	5	5%
Too Low?	47	45%
Total	104	100%

<i>KCCT Reading</i>		
About Right?	83	81%
Too High?	7	7%
Too Low?	13	13%
Total	103	100%
<i>KCCT Mathematics</i>		
About Right?	82	80%
Too High?	7	7%
Too Low?	14	14%
Total	103	100%
<i>KCCT Science</i>		
About Right?	78	76%
Too High?	18	17%
Too Low?	7	7%
Total	103	100%
<i>KCCT Social Studies</i>		
About Right?	78	76%
Too High?	17	17%
Too Low?	8	8%
Total	103	100%
<i>KCCT Arts and Humanities</i>		
About Right?	53	51%
Too High?	43	42%
Too Low?	7	7%
Total	103	100%
<i>KCCT Practical Living/Vocational Skills</i>		
About Right?	57	55%
Too High?	40	38%
Too Low?	7	7%
Total	104	100%
Nonacademic Factors		
<i>Attendance Rate</i>		
About Right?	74	71%
Too High?	16	15%
Too Low?	14	13%
Total	104	100%
<i>Retention Rate</i>		
About Right?	71	70%
Too High?	23	23%
Too Low?	8	8%
Total	102	100%

<i>Dropout Rate</i>		
About Right?	65	64%
Too High?	26	25%
Too Low?	11	11%
Total	102	100%

<i>Transition to Adult Life</i>		
About Right?	65	64%
Too High?	30	30%
Too Low?	6	6%
Total	101	100%

<i>16: Should any of the above components be removed?</i>		
Yes	49	48%
No	38	37%
Not Sure	16	16%
Total	103	100%

<i>17: Should other subject areas or nonacademic factors be included in the assessment and accountability system?</i>		
Yes	11	11%
No	78	76%
Not Sure	13	13%
Total	102	100%

18: What changes, if any, would you make to the CATS assessments or the accountability system?

19: In terms of direct costs of CATS testing and assessment to your school district, what percentage of your time is spent on specific CATS testing and assessment tasks?

<i>20: Does your district have a District Assessment Coordinator (DAC)? If YES, what percentage of his or her time is devoted to specific CATS testing and assessment tasks?</i>		
Yes	103	99%
No	1	1%
Total	104	100%

<i>21: Do you think all of the schools in your district can reach proficiency by 2014?</i>		
Yes	44	43%
No	26	25%
Not Sure	32	31%
Total	102	100%

Questions 22-24 deal with the federal No Child Left Behind (NCLB) Act.

22: Do you think the schools in your district can close the following achievement gaps by 2014?

Gender Achievement Gap

Yes	78	75%
No	6	6%
Not Applicable	2	2%
Not Sure	18	17%
Total	104	100%

Disability Achievement Gap

Yes	29	28%
No	44	42%
Not Applicable	1	1%
Not Sure	30	29%
Total	104	100%

English Language Proficiency Achievement Gap

Yes	31	30%
No	17	16%
Not Applicable	35	34%
Not Sure	21	20%
Total	104	100%

Income Achievement Gap

Yes	44	43%
No	34	33%
Not Applicable	3	3%
Not Sure	22	21%
Total	103	100%

Race/Ethnicity Achievement Gap

Yes	50	48%
No	13	13%
Not Applicable	26	25%
Not Sure	15	14%
Total	104	100%

23: Do you think your district can have a "highly qualified" teacher in every classroom by 2005?

Yes	41	40%
No	33	32%
Not Sure	29	28%
Total	103	100%

24: Are there any provisions of the national No Child Left Behind Act that you think will be difficult for schools in your district to achieve?

Yes	51	50%
No	19	18%
Not Sure	33	32%
Total	103	100%

If YES, which provision(s) of No Child Left Behind do you think will be most difficult to achieve?

Appendix E

Calculating Correlations Between ACT and CATS Scores

To correlate students' CATS scores with ACT scores, it was necessary to first build a data set with student-level scores for each of the eight academic components that are used to calculate a school's assessment index. For a student to have a score for each of these eight components, he or she must have been tested in each of grades 10, 11, and 12. CATS scores were available for 1999 through 2002; therefore, only those students who began 10th grade in 1999 or 2000 would have scores for all three years of testing. These student-level scores were then matched to individual ACT scores.

Grades and Tested Subjects, 1999 to 2002

Grade 10 (1999 or 2000)	Grade 11 (2000 or 2001)	Grade 12 (2001 or 2002)
Reading, Practical Living	Math, Science, Social Studies, Arts and Humanities	On-demand Writing, Writing Portfolio

Source: Kentucky Department of Education.

The Kentucky Department of Education provided the ACT, Kentucky Core Content Test (KCCT) scores, and on-demand writing and writing portfolio scores used for these analyses. The ACT data consisted of yearly scores from 1999 through 2002. Each ACT file contained the students' identifying information and individual-scaled scores for math, reading, and science reasoning, and a composite score. CATS scores consisted of KCCT and writing assessment data in individual files from 10th, 11th, and 12th grade for each year since 1999, the first year CATS testing was administered. The 10th-grade student files contained descriptive information on the student, as well as raw and scaled scores for the reading and practical living components of the KCCT. The 11th-grade files contained the same student information and raw and scaled scores for the math, science, social studies, and arts and humanities components of the KCCT. The 12th-grade files contained the same student information and final scores for the on-demand and portfolio writing sections of the CATS assessment.

After the students' scores were matched for 10th, 11th, and 12th grade, the scaled scores were used to calculate a student's composite score for the CATS assessment using the high school academic index formula without the CTBS score and the nonacademic factors.

Of the student files received from KDE, approximately 59,000 had scores from 10th, 11th, and 12th grades and therefore, had scores for all academic components of the CATS assessment. ACT scores were received for approximately 96,800 students who took the test from 1999 through 2002 in Kentucky. After matching student-level CATS

assessment scores with ACT scores, 41,399 students, or 70 percent of those students with three years of CATS scores, remained in the data set.

Appendix F

Statement of NTAPAA on the Validity and Reliability of CATS

The Kentucky Commonwealth Accountability Testing System: Appraising the System's Validity August 2003

National Technical Advisory Panel on Accountability and Assessment (NTAPAA)

James S. Catterall, UCLA, Chair
John Poggio, University of Kansas, Vice Chair
Suzanne Lane, University of Pittsburgh
Robert Linn, University of Colorado
David Miller, University of Florida
Andrew Porter, University of Wisconsin

Prepared for the Kentucky Legislative Research Commission

Introduction

This document is a response to a request from the Kentucky Legislative Research Commission to provide a statement articulating NTAPAA's assessment of the "validity" of Kentucky's education accountability and assessment system.

We begin with an outline of guiding principles or standards for state education accountability and assessment systems, followed by expressions of NTAPAA's appraisal of CATS' alignment with these principles. These standards, implicitly and explicitly, have guided NTAPAA in its reviews of KCCT and CATS. The standards are drawn from the experiences of panel members with the design and implementation of state accountability and assessment systems. The panel's work generally adheres to guidelines for testing and measurement catalogued in *Standards for Educational and Psychological Testing*, a work published jointly in 1999 by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education.

We note that the main concerns addressed in this document reflect the issues raised in two external reviews of the prior system, the Kentucky Instructional Results Information System.¹ These reports drew attention to the importance of periodic

¹ Hambleton, R. K., Jaeger, R. M., Koretz, D., Linn, R. L., Millman, J., and Phillips, S. E. (1995). *Review of the Measurement Quality of the Kentucky Instructional Results Information System, 1991-1994*. Frankfort: Office of Education Accountability, Kentucky General Assembly, June.

J. S. Catterall, W. Mehrens, J. Ryan, G. Flores, and P. Rubin (1998). "The Kentucky Instructional Results Information System: A Technical Review." Frankfort, KY: KY Legislative Research Commission, (January).

standard setting, linking and equating tests, assessing comparisons between system indicators and external measures, procedural quality controls, and the documentation of procedures and data used in developing and monitoring the performance of the components of the assessment and accountability system.

I. Student achievement assessments should mirror established State, LEA, and school instructional goals.

Instructional goals for state education systems are not carved in some historic stone, despite the evident similarities across the nation's education systems. Instructional goals should be clearly articulated in an assessment and accountability system. Goals should be established through processes appropriately involving system stakeholders.²

Kentucky works from a set of publicly developed academic expectations that guided the development of its Core Content for Assessment. This Core Content in turn is the basis for the development of tests addressing specific content areas at specific grade levels. CATS achievement assessments are well aligned with Kentucky's stated instructional goals.

Through core content tests and writing portfolios CATS gauges student achievement in multiple academic content areas, the arts and humanities, practical living and vocational skills. The system also incorporates measures of graduation rates and student retention. All measures contribute to a school index. Kentucky's school index scores lie at the heart of the accountability component of the system. Index scores for all schools must meet biennial growth targets leading to "proficiency" of all students by 2014.

NTAPAA members concur that the breadth of instructional content embraced by CATS contributes positively to its effectiveness." This breadth encourages teachers and schools to excel across the curriculum.

II. Standards for judging the nature and extent of student performance should be developed carefully and with participation of stakeholders.

Kentucky recently undertook a very broad-based process of re-setting standards. The process involved more than 2000 individuals. Two longstanding methods and one recently-developed method were employed for setting standards -- determining what constitutes proficient and other levels of student performance in each of the content areas and grade levels covered by Kentucky's core content tests. Participants discussed and appraised actual student work to comprehend ranges of student performance and to ground discussions of what student performance levels should be considered as novice, apprentice, proficient, and distinguished. Participants reviewed sample

² Stakeholders in this sense include but should not be limited to professional educators and system staff members, parents, community members including employers, and appointed or elected educational advisory boards.

questions to produce consensus on what sorts of questions students should be capable of answering if they were to be considered to be performing at various designated achievement levels. Members of NTAPAA attended parts of the standard setting activities.

Kentucky's standards setting process for CATS was as thorough and "state of the art" as any we are aware of. The standards reflect the judgments of a broad base of educators and citizens of Kentucky. Students judged to be doing proficient work through the measures produced by CATS should be considered on average to be accomplished and competent – in the spirit and meaning of proficiency implied by the Kentucky Education Reform Act and goals currently set by the Kentucky Board of Education for all schools.

II. Standards should be revisited periodically to check on their alignment with stakeholder expectations and values concerning formal gradations of student proficiency.

As just described, Kentucky re-set its standards in 1999 and expects to revisit standards again at a suitable interval.

III. An accountability system should be built on a platform of valid and reliable measures of achievement and/or behaviors of system participants.

Kentucky Department of Education staff, assessment contractors, and long-term consultants (e.g. HumRRO), along with NTAPAA and advisory groups such as the School Curriculum Advisory Committee, engage in ongoing monitoring, discussion, and formal studies of CATS' component measures and system practices. These activities include, but are not limited to,

- A. Generating and vetting test items.*
- B. Reviewing the psychometric properties of test items and their resulting suitability for inclusion in operational tests.*
- C. Reviewing test items for bias.*
- D. Designing test forms (e.g. six alternate forms of a given test).*
- E. Equating multiple forms of tests.*
- F. Designing scoring rubrics for open-response items consistent with established standards.*
- G. Monitoring scoring and scoring related activities (e.g. reviewing and observing scorer training and actual scoring activities).*
- H. Annual auditing of local scoring of writing portfolios.*
- I. Investigating and exploring differences in KCCT student performance by SES, gender, ethnicity, English Proficiency status, and disabilities.*
- J. Deliberating the content and format of score reports and school report cards.*
- K. Seeking external sources of validation which can help anchor the meaning of and attributions for content scores and school index changes in CATS. In*

other words, how are KCCT scores related to other widely obtained measures of educational achievement?

- L. Investigating the degree to which changes in CATS scores at the school level are mirrored by respective improvements in instructional strategies.*
- M. Projecting psychometric properties of discussed or proposed additions or changes in CATS, such as the use and augmentation of norm-referenced tests to satisfy testing in all grades under NCLB.*

KDE staff provide an ongoing flow of information for discussion and appraisals by NTAPAA. KDE initiates panel discussions of issues and brings relevant data to these discussions. KDE frequently, and with dexterity and timeliness, responds to requests for NTAPAA for differing or additional data and for alternative ways of presenting or formatting data. A consensus among NTAPAA members is that KDE maintains a more than adequate level and scope of documentation concerning issues and practices related to CATS. KDE has been forthcoming with information relative to the panel's technical concerns, and the KDE, its contractors, and CATS as a system have been very responsive to concerns and suggestions of the technical panel.

It must be said that the listing, A. through M. above, embraces many considerations, sub-issues, and sub-questions, along with multiple approaches to considering the merits of specific components or procedures within KCCT and CATS. Some of these inquiries have gone forward and have yielded results satisfactory to NTAPAA. Many are the subjects of ongoing or periodic deliberations. Some are planned inquiries.

As far as NTAPAA is concerned, KDE has amply considered the priorities brought to discussions of and inquiry into issues of validity and reliability of KCCT and CATS -- questions have been raised and addressed in an appropriate order given the demands of a developing system. Some inquiries, such as certain external validations of score changes or index growth, can only be done effectively with multiple years of data. These types of studies are coming on line at the present time or have been carried out in preliminary ways.

- V. An accountability system should recognize that there is error caused by a number of factors attaching to all component measures. In the case of summative judgments about whether schools or other units have met their growth achievement targets, a system benefits from recognizing margins for error in classifying schools as meeting or not meeting targets.**

CATS provides an error band approximating the standard error of measurement around calculated school achievement indexes. Thus schools attaining an index within one standard error of measurement of their targets are considered to have met index targets. This reduces the probability that schools which actually met their goals will be incorrectly designated as not having met goals due to aggregate measurement error.

VI. Summary Statement

There is substantial evidence supporting the validity and reliability of the Kentucky Core Content Test and the validity of the Commonwealth Accountability Testing System in the context of current implementation and use in the State of Kentucky. The information generated through this system provides valid and reliable indicators of student performance across schools and districts. Statistically significant changes in index scores over time should be considered valid indications that average student performance levels have changed over time. The evidence of student performance produced through KCCT and CATS constitutes a valid basis for rewarding or identifying for improvement schools and school districts to the degree provided in current statutes and regulations.

No Child Left Behind Act Standards and Assessment Provisions

STATE	Reading Standards	Math Standards	Science Standards	Annual Assessments in Reading	Annual Assessments in Math	Assessments in Science	Assessments in English Language Proficiency (LEP)	Inclusion of LEP Students	Inclusion of Students With Disabilities	Inclusion of Migrant Students	Disaggregation of Results
Alabama	Y	Y	Y	P	P	P	N	N	Y	N	Y
Alaska	P	P	P	P	P	N	Y	Y	Y	P	N
Arizona	Y	Y	Y	P	P	N	Y	Y	Y	Y	P
Arkansas	Y	Y	Y	P	P	N	Y	P	Y	Y	N
California	Y	Y	Y	Y	Y	Y	Y	P	Y	P	P
Colorado	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
Connecticut	P	P	Y	P	P	N	Y	P	Y	P	Y
Delaware	P	P	Y	Y	Y	Y	Y	Y	Y	Y	Y
District of Columbia	Y	Y	Y	P	P	N	Y	P	P	P	N
Florida	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	P
Georgia	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	P
Hawaii	P	P	Y	Y	Y	P	Y	P	Y	Y	Y
Idaho	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
Illinois	Y	Y	Y	P	P	Y	Y	Y	Y	P	P
Indiana	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Iowa	Y	Y	Y	P	P	P	Y	P	Y	Y	Y
Kansas	P	P	Y	P	P	P	Y	Y	Y	Y	P
Kentucky	P	P	Y	P	P	Y	N	N	Y	N	Y
Louisiana	Y	P	Y	P	P	Y	N	Y	Y	Y	P
Maine	P	P	Y	P	P	N	P	Y	Y	Y	N
Maryland	P	P	Y	Y	Y	Y	Y	Y	Y	Y	Y
Massachusetts	P	P	Y	P	P	P	Y	Y	Y	Y	P
Michigan	P	P	Y	P	P	Y	P	P	Y	Y	P
Minnesota	Y	Y	Y	P	P	N	Y	Y	Y	U	Y
Mississippi	Y	Y	Y	Y	Y	P	Y	P	Y	P	Y
Missouri	P	P	Y	P	P	Y	P	Y	Y	Y	N
Montana	P	P	Y	P	P	Y	N	Y	Y	Y	Y
Nebraska	P	P	Y	P	P	Y	Y	Y	Y	Y	P
Nevada	Y	Y	Y	P	P	Y	P	N	P	Y	Y
New Hampshire	P	P	P	P	P	P	Y	Y	Y	N	N
New Jersey	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	P
New Mexico	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

No Child Left Behind Act Standards and Assessment Provisions

STATE	Reading Standards	Math Standards	Science Standards	Annual Assessments in Reading	Annual Assessments in Math	Assessments in Science	Assessments in English Language Proficiency (LEP)	Inclusion of LEP Students	Inclusion of Students With Disabilities	Inclusion of Migrant Students	Disaggregation of Results
New York	P	P	P	P	P	Y	Y	Y	Y	Y	Y
North Carolina	Y	Y	Y	Y	Y	P	Y	Y	Y	Y	Y
North Dakota	P	P	Y	P	P	Y	Y	Y	Y	Y	P
Ohio	Y	Y	Y	P	P	P	P	N	Y	P	P
Oklahoma	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Oregon	P	P	Y	P	P	Y	N	N	N	P	P
Pennsylvania	P	P	Y	P	P	P	N	Y	Y	Y	Y
Rhode Island	P	P	Y	P	P	N	Y	Y	Y	P	P
South Carolina	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	P
South Dakota	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	N
Tennessee	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Texas	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Utah	Y	Y	Y	Y	Y	U	P	Y	Y	Y	Y
Vermont	P	P	Y	P	P	Y	N	P	P	P	Y
Virginia	Y	Y	Y	P	P	Y	Y	Y	Y	Y	P
Washington	P	Y	Y	P	P	Y	Y	Y	Y	Y	Y
West Virginia	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y
Wisconsin	P	P	Y	P	P	Y	Y	Y	Y	Y	Y
Wyoming	P	P	Y	P	P	U	N	Y	Y	Y	Y
# of States on Track	28	28	48	20	20	30	35	36	45	37	26
% of States on Track	54%	54%	92%	38%	38%	58%	67%	60%	87%	71%	50%

Y=Appears to be on track, P=Appears to be partially on track, N=Does not appear to be on track, U=Data unavailable.

Source: Compiled by Program Review staff from the Education Commission of the States database.

<http://nclb.ecs.org/nclb> (July 29, 2003)

No Child Left Behind Act Accountability Provisions

STATE	Single Accountability System	All Schools and Students Included	Continuous Growth to 100% Proficiency	Annual Determination of Adequate Yearly Progress	Accountability for All Subgroups	Primarily Based on Academics	Includes Graduation Rate and Additional Indicator	Based on Separate Math and Reading Objectives	95% of Students in All Subgroups Assessed	State Report Cards
Alabama	Y	Y	N	N	N	N	Y	N	N	P
Alaska	N	N	N	N	U	P	Y	P	Y	P
Arizona	P	P	P	Y	P	Y	Y	Y	N	P
Arkansas	Y	Y	N	Y	Y	P	Y	Y	N	P
California	Y	Y	Y	Y	Y	Y	Y	P	Y	P
Colorado	N	N	P	Y	N	Y	Y	Y	P	P
Connecticut	Y	Y	N	Y	N	P	N	N	N	P
Delaware	Y	P	N	Y	N	Y	N	Y	N	P
District of Columbia	N	N	N	N	N	N	N	N	N	N
Florida	Y	Y	P	Y	P	Y	N	Y	P	P
Georgia	N	N	P	Y	P	Y	N	Y	P	P
Hawaii	N	N	N	Y	N	Y	P	Y	N	P
Idaho	N	Y	P	P	P	P	P	P	N	P
Illinois	N	N	Y	Y	N	N	Y	N	N	P
Indiana	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Iowa	Y	Y	P	P	Y	P	Y	Y	Y	P
Kansas	Y	Y	P	P	Y	Y	Y	Y	Y	Y
Kentucky	Y	P	P	P	P	Y	P	N	P	Y
Louisiana	Y	Y	P	P	N	Y	P	P	P	P
Maine	Y	Y	N	P	N	P	N	N	N	P
Maryland	Y	Y	P	Y	Y	P	P	Y	P	P
Massachusetts	Y	P	Y	Y	U	Y	P	Y	Y	P
Michigan	Y	Y	P	Y	Y	Y	P	Y	Y	P
Minnesota	Y	Y	N	Y	U	P	Y	P	N	P
Mississippi	Y	Y	Y	Y	N	Y	Y	Y	P	P
Missouri	Y	Y	P	Y	P	Y	P	Y	Y	P
Montana	U	N	N	N	N	N	N	N	N	P
Nebraska	P	P	N	P	P	P	N	N	N	P
Nevada	P	P	N	Y	N	Y	Y	P	N	Y

No Child Left Behind Act Accountability Provisions

STATE	Single Accountability System	All Schools and Students Included	Continuous Growth to 100% Proficiency	Annual Determination of Adequate Yearly Progress	Accountability for All Subgroups	Primarily Based on Academics	Includes Graduation Rate and Additional Indicator	Based on Separate Math and Reading Objectives	95% of Students in All Subgroups Assessed	State Report Cards
New Hampshire	U	P	N	N	N	P	Y	P	N	P
New Jersey	Y	Y	P	P	P	Y	P	Y	Y	P
New Mexico	Y	Y	P	Y	N	Y	Y	Y	P	P
New York	Y	Y	Y	Y	P	Y	Y	Y	Y	P
North Carolina	Y	Y	Y	Y	Y	Y	P	Y	Y	Y
North Dakota	N	Y	N	P	N	Y	N	Y	N	P
Ohio	Y	Y	Y	Y	Y	Y	Y	Y	Y	P
Oklahoma	Y	Y	Y	Y	Y	Y	Y	Y	Y	P
Oregon	P	N	N	P	N	U	P	N	N	P
Pennsylvania	Y	Y	Y	Y	P	Y	Y	P	N	Y
Rhode Island	P	P	P	P	P	Y	P	P	Y	P
South Carolina	Y	Y	Y	Y	Y	Y	Y	Y	Y	P
South Dakota	Y	Y	Y	Y	Y	Y	Y	Y	P	P
Tennessee	Y	Y	P	Y	Y	Y	Y	P	P	P
Texas	N	N	P	Y	P	Y	Y	Y	P	P
Utah	Y	Y	N	P	N	Y	N	N	N	P
Vermont	Y	Y	Y	Y	Y	P	P	P	P	P
Virginia	P	Y	Y	Y	Y	Y	Y	Y	Y	P
Washington	Y	Y	P	P	P	P	P	P	N	Y
West Virginia	Y	Y	Y	Y	Y	Y	Y	P	Y	P
Wisconsin	Y	Y	P	Y	P	Y	N	Y	P	P
Wyoming	Y	Y	Y	Y	Y	Y	N	Y	Y	P
# of States on Track	34	34	16	33	17	34	25	27	18	7
% of States on Track	65%	65%	31%	63%	33%	65%	48%	52%	35%	13%

Y=Appears to be on track, P=Appears to be partially on track, N=Does not appear to be on track, U=Data unavailable.

Source: Compiled by Program Review staff from the Education Commission of the States database.

<http://nclb.ecs.org/nclb> (July 29, 2003)

**No Child Left Behind
School Improvement Provisions**

STATE	Timely Identification	Technical Assistance	Public School Choice	Rewards and Sanctions	School Recognition	School Restructuring	Corrective Action for Local Educational Agencies
Alabama	P	P	Y	P	N	Y	Y
Alaska	P	P	N	Y	Y	N	N
Arizona	P	P	Y	P	N	Y	N
Arkansas	Y	P	Y	Y	Y	Y	Y
California	P	P	Y	Y	Y	Y	Y
Colorado	P	P	Y	Y	Y	Y	Y
Connecticut	N	P	P	Y	Y	Y	P
Delaware	P	P	Y	Y	Y	Y	N
District of Columbia	N	N	N	N	N	N	U
Florida	Y	P	Y	Y	Y	Y	N
Georgia	P	P	Y	Y	Y	Y	N
Hawaii	N	Y	Y	Y	Y	N	U
Idaho	N	P	P	N	N	N	N
Illinois	P	P	Y	Y	Y	Y	Y
Indiana	P	P	P	Y	Y	N	N
Iowa	P	Y	Y	P	N	Y	Y
Kansas	P	P	P	Y	Y	Y	Y
Kentucky	Y	P	Y	Y	Y	Y	Y
Louisiana	Y	Y	Y	Y	Y	Y	Y
Maine	P	P	P	N	N	N	N
Maryland	P	P	Y	Y	Y	Y	P
Massachusetts	P	P	Y	Y	Y	Y	Y
Michigan	P	P	P	Y	Y	Y	P
Minnesota	P	P	P	N	N	N	N
Mississippi	Y	P	P	Y	Y	Y	Y
Missouri	P	P	P	P	N	Y	Y
Montana	N	N	P	N	N	N	N
Nebraska	N	P	P	N	N	N	N

No Child Left Behind Act School Improvement Provisions

STATE	Timely Identification	Technical Assistance	Public School Choice	Rewards and Sanctions	School Recognition	School Restructuring	Corrective Action for Local Educational Agencies
Nevada	P	P	P	Y	Y	Y	N
New Hampshire	N	P	P	N	N	N	N
New Jersey	N	P	P	P	Y	N	Y
New Mexico	P	P	P	Y	Y	Y	Y
New York	P	P	P	P	N	Y	P
North Carolina	Y	P	N	Y	Y	Y	Y
North Dakota	N	N	P	N	N	N	N
Ohio	P	P	Y	P	N	Y	P
Oklahoma	P	P	Y	Y	Y	Y	Y
Oregon	N	P	P	P	N	N	N
Pennsylvania	N	N	P	Y	Y	Y	Y
Rhode Island	P	Y	P	P	N	Y	Y
South Carolina	P	P	P	Y	Y	Y	Y
South Dakota	P	P	Y	P	P	P	P
Tennessee	Y	P	Y	Y	Y	Y	Y
Texas	P	P	Y	Y	Y	Y	Y
Utah	P	P	Y	N	N	N	N
Vermont	P	P	P	P	N	Y	N
Virginia	P	Y	Y	Y	Y	U	N
Washington	P	P	Y	N	N	N	N
West Virginia	P	P	Y	Y	Y	Y	Y
Wisconsin	Y	P	Y	N	N	N	N
Wyoming	P	P	P	P	N	N	N
# of States on Track	8	5	25	28	29	32	22
% of States on Track	15%	10%	48%	54%	56%	62%	42%

Y=Appears to be on track, P=Appears to be partially on track, N=Does not appear to be on track, U=Data unavailable.

Source: Compiled by Program Review staff from the Education Commission of the States database.

<http://nclb.ecs.org/nclb> (July 29, 2003)

No Child Left Behind Act Safe Schools Provisions

STATE	Criteria for Unsafe Schools	Transfer Policy for Students in Unsafe Schools	Transfer Policy for Victims of Violent Crimes
Alabama	Y	Y	Y
Alaska	P	P	P
Arizona	P	P	P
Arkansas	Y	Y	Y
California	Y	Y	P
Colorado	P	P	P
Connecticut	Y	Y	Y
Delaware	Y	Y	Y
District of Columbia	N	N	N
Florida	Y	Y	Y
Georgia	P	P	P
Hawaii	Y	Y	Y
Idaho	Y	Y	Y
Illinois	Y	Y	Y
Indiana	Y	Y	Y
Iowa	Y	Y	Y
Kansas	Y	Y	Y
Kentucky	Y	Y	Y
Louisiana	P	P	P
Maine	P	P	P
Maryland	P	P	P
Massachusetts	Y	Y	Y
Michigan	Y	Y	Y
Minnesota	N	N	N
Mississippi	Y	Y	Y
Missouri	Y	Y	Y
Montana	Y	Y	Y
Nebraska	N	N	N
Nevada	Y	Y	Y
New Hampshire	Y	Y	Y
New Jersey	Y	Y	Y
New Mexico	N	N	N
New York	Y	Y	Y
North Carolina	Y	Y	Y
North Dakota	P	P	P
Ohio	Y	Y	Y
Oklahoma	Y	Y	Y
Oregon	N	N	N
Pennsylvania	Y	Y	Y
Rhode Island	N	N	N
South Carolina	Y	Y	Y
South Dakota	Y	Y	Y
Tennessee	P	P	P
Texas	N	N	N
Utah	Y	Y	Y

No Child Left Behind Act Safe Schools Provisions

STATE	Criteria for Unsafe Schools	Transfer Policy for Students in Unsafe Schools	Transfer Policy for Victims of Violent Crimes
Vermont	Y	Y	Y
Virginia	Y	Y	Y
Washington	N	N	N
West Virginia	Y	Y	Y
Wisconsin	P	P	P
Wyoming	N	N	N
# of States on Track	32	32	31
% of States on Track	62%	62%	60%

Y=Appears to be on track, P=Appears to be partially on track, N=Does not appear to be on track, U=Data unavailable.

Source: Compiled by Program Review staff from the Education Commission of the States database.

<http://nclb.ecs.org/nclb> (July 29, 2003)

No Child Left Behind Act Supplemental Services Provisions

STATE	Criteria for Supplemental Services	List of Approved Supplemental Services Providers	Monitoring of Supplemental Services Providers	Implementation of Supplemental Services
Alabama	Y	Y	Y	P
Alaska	Y	Y	Y	P
Arizona	Y	Y	Y	Y
Arkansas	Y	Y	Y	Y
California	Y	Y	Y	Y
Colorado	Y	Y	Y	P
Connecticut	Y	Y	Y	P
Delaware	Y	N	Y	P
District of Columbia	N	Y	N	N
Florida	N	N	N	N
Georgia	Y	Y	N	Y
Hawaii	Y	Y	N	Y
Idaho	Y	Y	N	N
Illinois	Y	Y	Y	Y
Indiana	Y	Y	Y	P
Iowa	Y	Y	Y	Y
Kansas	Y	Y	N	P
Kentucky	Y	Y	Y	Y
Louisiana	Y	Y	N	P
Maine	Y	N	N	N
Maryland	Y	Y	Y	P
Massachusetts	Y	Y	Y	Y
Michigan	Y	Y	Y	P
Minnesota	N	N	N	N
Mississippi	Y	Y	N	Y
Missouri	Y	Y	Y	P
Montana	Y	Y	Y	Y
Nebraska	Y	P	N	N
Nevada	Y	P	Y	P
New Hampshire	Y	Y	N	N
New Jersey	Y	Y	N	N
New Mexico	Y	Y	N	P
New York	Y	Y	Y	Y
North Carolina	Y	Y	Y	P
North Dakota	Y	Y	Y	P
Ohio	Y	Y	Y	P
Oklahoma	Y	Y	Y	Y
Oregon	Y	N	N	N
Pennsylvania	N	Y	N	N
Rhode Island	Y	Y	Y	P
South Carolina	Y	Y	N	N
South Dakota	Y	Y	Y	Y

No Child Left Behind Act Supplemental Services Provisions

STATE	Criteria for Supplemental Services	List of Approved Supplemental Services Providers	Monitoring of Supplemental Services Providers	Implementation of Supplemental Services
Tennessee	Y	Y	N	P
Texas	Y	Y	Y	Y
Utah	Y	Y	Y	P
Vermont	P	Y	N	N
Virginia	Y	Y	Y	Y
Washington	Y	Y	Y	Y
West Virginia	Y	N	N	N
Wisconsin	Y	Y	Y	P
Wyoming	N	N	N	N
# of States on Track	45	43	30	17
% of States on Track	87%	83%	58%	33%

Y=Appears to be on track, P=Appears to be partially on track, N=Does not appear to be on track,
 U=Data unavailable.

Source: Compiled by Program Review staff from the Education Commission of the States database.

<http://nclb.ecs.org/nclb> (July 29, 2003)

No Child Left Behind Act Teacher Quality Provisions

STATE	Highly Qualified Teachers Definition	Subject Matter Competence	Test for New Elementary Teachers	Highly Qualified Teacher in Every Classroom	Highly Qualified Professional Development
Alabama	Y	Y	Y	N	N
Alaska	N	N	N	N	N
Arizona	Y	P	Y	N	N
Arkansas	P	U	Y	N	N
California	P	P	Y	P	N
Colorado	N	Y	Y	N	N
Connecticut	Y	Y	Y	N	Y
Delaware	P	P	Y	N	N
District of Columbia	N	N	N	N	N
Florida	Y	Y	Y	N	N
Georgia	P	P	Y	N	N
Hawaii	P	P	P	N	N
Idaho	P	P	Y	N	N
Illinois	Y	P	Y	N	N
Indiana	P	P	Y	N	Y
Iowa	Y	Y	Y	N	P
Kansas	Y	Y	Y	N	P
Kentucky	P	Y	Y	N	N
Louisiana	N	N	P	N	N
Maine	N	P	Y	N	P
Maryland	Y	P	Y	N	N
Massachusetts	U	P	Y	N	N
Michigan	P	P	Y	N	N
Minnesota	P	P	Y	N	N
Mississippi	P	Y	Y	P	P
Missouri	N	P	Y	N	N
Montana	Y	N	Y	N	N
Nebraska	P	P	N	N	N
Nevada	P	P	Y	P	N
New Hampshire	N	P	Y	N	P
New Jersey	P	P	Y	P	U
New Mexico	N	P	Y	N	N
New York	N	P	Y	N	N
North Carolina	P	P	Y	N	N
North Dakota	N	P	Y	N	N
Ohio	P	N	Y	N	N
Oklahoma	P	P	Y	N	P
Oregon	P	P	Y	N	N
Pennsylvania	N	P	Y	P	N
Rhode Island	N	P	P	N	N
South Carolina	P	P	P	P	P
South Dakota	P	P	P	N	N

No Child Left Behind Act Teacher Quality Provisions

STATE	Highly Qualified Teachers Definition	Subject Matter Competence	Test for New Elementary Teachers	Highly Qualified Teacher in Every Classroom	Highly Qualified Professional Development
Tennessee	N	N	Y	N	N
Texas	N	Y	Y	N	P
Utah	N	N	N	N	N
Vermont	N	P	Y	N	N
Virginia	Y	Y	Y	P	N
Washington	N	P	Y	N	N
West Virginia	N	P	Y	N	N
Wisconsin	P	P	Y	Y	N
Wyoming	P	N	N	N	N
# of States on Track	10	10	41	1	2
% of States on Track	19%	19%	79%	2%	4%

Y=Appears to be on track, P=Appears to be partially on track, N=Does not appear to be on track, U=Data unavailable.

Source: Compiled by Program Review staff from the Education Commission of the States database.

<http://nclb.ecs.org/nclb> (July 29, 2003)

**Kentucky Department of Education Response
Program Review and Investigations Committee Draft Report
The Commonwealth Accountability Testing System
August 22, 2003**

RECOMMENDATIONS

2.1

“review the schools’ documentation that students coded as transfers are enrolled in other schools”

- With the implementation of the student tracking system, KDE will have the ability to verify student transfers within the state. We will not have the capacity to electronically track out-of-state transfers, except to note when a transcript is requested.
- KDE currently reviews student withdrawal information as part of the attendance audit process. Within this process, staff verifies the use of the appropriate withdrawal code in the STI software. For example, student transcript requests are reviewed for students who are coded as transfers. There has been no indication of widespread problems with misreporting.
- We will thoroughly review, make appropriate changes to the audit procedure, and update and simplify withdrawal codes.

2.2

“consider sanctioning schools that underreport statistics by lowering their scores on the accountability index by an additional amount or by making them ineligible for rewards that year.”

- Two sanctions already in place appear to control misreporting by schools and districts. When errors in the nonacademic data are identified, corrections to the Commonwealth Accountability Testing System data are made. Such corrections can result in changes in school performance judgments. An intentional error violates the “Administration Code” – 703 KAR 5:080. This calls for the district to consider disciplinary action and to report the violation to the Education Professional Standards Board.
- KDE wants to pursue 2.1 immediately and report results before 2.2 actions are considered.

2.3

“Implement a uniform student information system at the state level to track students who transfer, drop out, are expelled, or graduate; and to keep track of the number of students who are excluded from the CATS test.”

- A statewide system is being implemented. In December, KDE will field test the secure student identifier system and anticipates implementation of the state data system in the 2004-2005 school year. KDE must test the system for accuracy and security and collect data for students over time before the system can be fully utilized to track students for this purpose.

2.4

“create a CATS testing expenditure category in the MUNIS system and encourage school districts to utilize this category for all CATS administrative expenses”

- KDE will create a CATS expenditure sub-code in the MUNIS system. Getting uniform and consistent information is more difficult, will be time consuming for schools and districts, and require training to support. Instruction without good measurement will not lead to improved student learning. Assessment must be seen as part of an instructional and professional development system to better meet the needs of students.
- We are working with several districts to determine the amount of time being spent on assessment activities.

4.1

“Reduce . . . the practice of teachers scoring their own students’ portfolios.”

- One of our current options, double-blind scoring, accomplishes two purposes: it provides a means for the teacher of record to assess and assist students and it helps to alleviate the impact of scorer bias because no one scorer is involved in the process. Scores of record are determined by agreement. This study indicates that approximately 85% of Kentucky schools already use this option (LRC study p.61). KDE will exert more guidance to promote sound models.
- Through the KDE Cluster Leader Electronic Messaging system, KDE will survey cluster leaders to determine methods used by schools to alleviate the concern of a teacher scoring his/her own students’ writing portfolios. This information will then be shared with all cluster leaders via the electronic system.

“survey teachers to determine how their portfolio scoring training can be improved”

- Using the electronic messaging system, KDE will survey teachers to gather suggested improvements to portfolio scoring training. These suggested changes will be taken under advisement by the Kentucky Writing Advisory Committee and the Kentucky Board of Education.

“regularly replace benchmark portfolios with new samples”

- Maintaining the benchmarks helps Kentucky teachers understand writing standards. The repetition helps scorers anchor the language of the performance cells on the scoring guide. Current benchmarks will need to remain a consistent component of the portfolio assessment.
- Moreover, additional portfolios will be released every year to assist with scoring and instruction.

“provide teachers with more opportunities to practice scoring writing portfolios”

- KDE will provide quality control portfolios for scoring practice.
- KDE will offer scoring verification sessions throughout the state for teachers to determine their own scoring accuracy.

- Using the electronic messaging system, KDE will suggest methods for cluster leaders to offer scoring practice throughout the year (e.g., use of quality control portfolios currently on file, use of audited portfolios).
- As a result of the loss of the Regional Service Centers, including the regional writing consultants, our capacity to provide additional face-to-face opportunities for this practice is greatly diminished. We are exploring the possibility of electronic practice opportunities.

4.2

“track the performance of portfolio scorers”

- KDE will review the use of record-keeping forms with all cluster leaders.
- KDE will develop questions to assist with scoring accuracy research. These questions will be distributed to cluster leaders throughout the state; results will be collected and reviewed by KDE staff. Significant findings will be presented to the Kentucky Writing Advisory Committee. Suggested modifications to scoring practices will be shared with all cluster leaders.

4.3

“establish consequences for schools that have low portfolio audit agreement rates”

- Applying the “final audit” score is often considered a serious consequence by audited schools and teachers typically view the audit experience as embarrassing. If the “inaccurate scoring” is intentional or appropriate training was not provided, the “Administration Code” provisions apply.
- Although writing portfolio scores may be adjusted during the writing portfolio audit, KDE views the audit as a means to support schools and inform them of their instructional and assessment needs.
- Consequently, KDE will postpone decisions on punitive measures until all other areas have been addressed and schools have had an opportunity to respond.

“consider re-auditing schools that had a high number of scoring inaccuracies the prior year”

- KDE will discuss the recommendation with the National Technical Advisory Panel on Assessment and Accountability.
- Based on the selection procedure, schools can be audited in consecutive years. The data indicate that schools often improve writing instruction and scoring accuracy during the second audited year.

“consider increasing the number of schools randomly selected for audits”

- As of the 2002-2003 Writing Portfolio Audit, 80% of audited schools will be randomly selected. (90 of the 113 audited schools were randomly selected during the 2003 audit.)

Program Review and Investigations Committee

LEGISLATIVE RESEARCH COMMISSION

FRANKFORT, KENTUCKY

Printed with state funds.