

University of Michigan Journal of Law Reform

Volume 28

1995

Achieving Equity and Excellence in Kentucky Education

C. Scott Trimble

Kentucky Department of Education

Andrew C. Forsaith

Follow this and additional works at: <https://repository.law.umich.edu/mjlr>



Part of the [Constitutional Law Commons](#), [Education Law Commons](#), [Litigation Commons](#), and the [State and Local Government Law Commons](#)

Recommended Citation

C. S. Trimble & Andrew C. Forsaith, *Achieving Equity and Excellence in Kentucky Education*, 28 U. MICH. J. L. REFORM 599 (1995).

Available at: <https://repository.law.umich.edu/mjlr/vol28/iss3/6>

This Symposium Article is brought to you for free and open access by the University of Michigan Journal of Law Reform at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in University of Michigan Journal of Law Reform by an authorized editor of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.

ACHIEVING EQUITY AND EXCELLENCE IN KENTUCKY EDUCATION

C. Scott Trimble*
Andrew C. Forsaith**

In this Article, Trimble and Forsaith discuss the landmark Kentucky school finance case, Rose v. Council for Better Education, 790 S.W.2d 186 (Ky. 1989), and the school reform efforts it spawned. In Council for Better Education the Kentucky Supreme Court held that the state had failed its duty under the state constitution to provide all students with an adequate education, which it defined in terms of seven categories of knowledge and skills students should acquire. The State General Assembly responded with the Kentucky Education Reform Act (KERA), which significantly boosted state funding as well as established an ambitious accountability system based on high academic standards for all students. Trimble and Forsaith explore the extent to which Council for Better Education and KERA mark a major departure from previously modest reform efforts in Kentucky and attempt to answer what brought them about. In addition, the authors discuss the substantial challenges involved in implementing the Council for Better Education mandate by examining the central component of KERA, the new statewide assessment system.

Kentucky embarked upon massive school reform with the passage of the Kentucky Education Reform Act (KERA) in 1990.¹ Nowhere had such comprehensive reform been attempted on a statewide level. KERA mandated school-based decision making, statewide curriculum frameworks, and an ambitious accountability system with rewards and sanctions for schools tied to the achievement of high academic standards for all students. In terms of school finance KERA established the Support Education Excellence in Kentucky (SEEK) fund,² a new

* Director, Division of Assessment Implementation, Office of Curriculum Assessment and Accountability, Kentucky Department of Education. B.A. 1971, M.A. 1972, Michigan State University. After Mr. Trimble wrote the part of this Article concerning the implementation of the Kentucky Education Reform Act (KERA), the Kentucky Department of Education issued a number of manuals on KERA. References to these manuals have been included in order to provide the reader with additional resources.

** B.A. 1988, University of Chicago; J.D. 1994, University of Wisconsin.

1. Kentucky Education Reform Act of 1990, ch. 476, 1990 Ky. Acts 1208 (codified as amended in scattered sections of KY. REV. STAT. ANN., chs. 156-65 and other scattered chapters) [hereinafter KERA].

2. KY. REV. STAT. ANN. §§ 157.310-.440 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

foundation program which increased the guaranteed minimum per-pupil expenditure statewide. This new program, combined with increased local spending, meant that Kentucky spent \$700 million more on its schools in 1991–1992 than in 1989–1990.³

The precipitating event for the Kentucky school reform was the Kentucky Supreme Court's decision in *Rose v. Council for Better Education*,⁴ which declared the entire state school system unconstitutional under Kentucky's constitution. In *Council for Better Education*, the court ruled that the General Assembly failed to perform its constitutional duty to provide an "efficient system of common schools" because of both the overall inadequacy of education in Kentucky and the great disparity in educational opportunity across the state.⁵ To remedy these problems, the court mandated that the State provide a school system that is substantially uniform across the state and sufficiently funded to ensure that every child receive an adequate education.⁶ Most notably, the court defined "adequate education" in terms of seven "capacities" which the court concluded all students should acquire.⁷

Part I of this Article describes the extent to which *Council for Better Education* and KERA mark a significant departure from Kentucky's past system of public education and presents some theories for why the change occurred. Despite the fact that concern for quality and equity in public education has been present throughout Kentucky's history, *Council for Better Education* and KERA demonstrate a quantum leap from previous reform efforts. Part II discusses the substantial challenges that the state faces in translating *Council for Better Education's* mandate into improvement in thousands of classrooms across the state by describing the new statewide assessment system, the central mechanism for implementing the court's mandate. The assessment system is charged with transforming the broad goals set forth by *Council for Better Education* into specific targets for school performance, for directing instructional improvement, and for monitoring the overall progress of the reform. Assessment encompasses most, if not all, of the challenges involved in creating a high performance school system.

3. 1992 KY. OFFICE OF EDUC. ACCOUNTABILITY ANN. REP. 40 [hereinafter KY. OFFICE OF EDUC. ACCOUNTABILITY].

4. 790 S.W.2d 186 (Ky. 1989).

5. *Id.* at 205–09.

6. *Id.* at 211–12.

7. *Id.* at 212.

I. THE HISTORY OF EDUCATION REFORM IN KENTUCKY

A. Early Equalization Measures

Kentucky historically has provided support to its public schools at a level well below the national average. At the time of the *Council for Better Education* decision, Kentucky ranked forty-third in the nation in expenditures per pupil and thirty-seventh in average teacher's salary.⁸ Performance was also deficient in many ways. For example, only 68.2% of ninth graders in Kentucky finished high school.⁹ Education was especially poor in Kentucky's Appalachian counties, where 48.4% of the population was functionally illiterate.¹⁰ In seventy-eight counties, more than half of the adults lacked a high school diploma.¹¹

Prior to *Council for Better Education* and KERA, Kentucky did attempt to improve its educational system's quality and financial equity. For years, state financial support had been limited by section 186 of the state constitution, which required state education funds to be distributed to school districts according to each district's school-age population, regardless of its wealth.¹² In the 1940s, section 186 was amended a number of times to eventually allow ten percent of state school funds to be used to equalize funding among districts.¹³ In 1953, the constitution was amended to eliminate the requirement of per capita expenditures.¹⁴

Exercising its new authority in 1954, Kentucky's General Assembly passed the state's first school funding equalization program, known as the Minimum Foundation Program (MFP).¹⁵ The MFP guaranteed a minimum educational expenditure for

8. *Rose v. Council for Better Educ., Inc.*, 790 S.W.2d 186, 197 (Ky. 1989).

9. *Id.*

10. Anne E. Keating, *You Get What You Pay For: Financing Public Schools in Kentucky*, KY. BENCH & BAR, Winter 1993, at 6 (quoting Brief for Appellees at 1-2, *Council for Better Educ.* (No. 88-SC-804-TG)).

11. *Id.*

12. KY. CONST. § 186 (amended 1953).

13. *See Act of Mar. 18, 1940, ch. 64, 1940 Ky. Acts 294* (repealed 1953).

14. *Act of Mar. 24, 1952, ch. 89, 1952 Ky. Acts 245* (approved 1953) (amending KY. CONST. § 186).

15. *Act of Mar. 22, 1954, ch. 214, 1954 Ky. Acts 590* (repealed 1974). The General Assembly had established an equalization fund in 1930, but it was struck down by the Kentucky Supreme Court in *Talbot v. Kentucky State Bd. of Educ.*, 52 S.W.2d 727 (Ky. 1932).

each district based on the number of "classroom units" in each district.¹⁶ The classroom units were allotted according to the number of students in the district. Extra units were added for vocational and special education services, administrative and instructional staff, and transportation services that were provided by the district.¹⁷ The state provided funds equal to the difference between the state guaranteed expenditure and the required local property tax effort of \$1.10 per \$100 of assessed value.¹⁸

Although the MFP offered some relief to poor districts, inequities remained because of the wide variance in assessed property values across the state. On average, property was assessed at only 27% of its fair market value, but it was also assessed as low as 12.5% in some communities and as high as 33% in others.¹⁹ The result was not only limited revenue for local districts, but also unequal revenues.

A suit was filed in the early 1960s arguing that these property assessment practices violated the state constitution. In *Russman v. Luckett*,²⁰ the Kentucky Supreme Court agreed, declaring that section 172 of the state constitution²¹ required that all property be assessed at 100% of fair market value.²² This decision would have more than tripled property taxes, on average, across the state.²³ However, legislative response was swift. A special session of the General Assembly enacted the so-called "Rollback Law,"²⁴ which reduced school, county, and city property tax levies to the 1965 level, except for revenue growth resulting from new property.²⁵ In a concession to school dis-

16. 1954 Ky. Acts 594-96.

17. *Id.* at 594, 596.

18. *See id.* at 597.

19. *Rose v. Council for Better Educ., Inc.*, 790 S.W.2d 186, 194 (Ky. 1989).

20. 391 S.W.2d 694, 700 (Ky. 1965).

21. Section 172 of the Kentucky Constitution reads:

All property, not exempted from taxation by this Constitution, shall be assessed for taxation at its fair cash value, estimated at the price it would bring at a fair voluntary sale

Id. at 696 (quoting KY. CONST. § 172).

22. *Id.*

23. *See supra* text accompanying note 19.

24. Act of Sept. 16, 1965, ch. 7, § 2, 1965 Ky. Acts 3, 6-9 (codified at KY. REV. STAT. ANN. § 160.470 (Michie/Bobbs-Merrill 1994)).

25. *Id.*

tricts, the Rollback Law allowed boards of education to make two one-time, ten percent increases in their tax levies in 1967 and 1968.²⁶

Thus, the new legislation reinstated the inequalities in school finance that *Russman* had momentarily eliminated. The situation arguably worsened over the following years when the Kentucky legislature allowed districts to levy one of three specialized taxes: (1) an occupational tax on wages and profits; (2) a tax on gross utility receipts; and (3) an excise tax on income.²⁷ Even cursory examination reveals that these taxes favored wealthy, heavily populated districts.

The last pre-Council for Better Education effort at funding equity was the Power Equalization Program (PEP), passed in 1976.²⁸ This program provided only minimal aid to local districts; 9¢ per \$100 of assessed value in 1985–1986 and 13¢ per \$100 of assessed value in 1986–1987.²⁹ Even this modest support was undercut by the General Assembly in 1979,³⁰ in the midst of the anti-tax fervor that swept the nation after California's Proposition 13.³¹ The new legislation required districts to decrease their tax rates so that the current year's revenue would not exceed the previous year's by more than four percent.³² Consequently, tax rates dropped on average from 31.5¢ to 22.9¢ per \$100 of assessed value as real property values rose around the state.³³

In the words of the Kentucky Supreme Court, "[i]f one were to summarize the history of school funding in Kentucky, one might well say that every forward step taken to provide funds to local districts and to equalize money spent for the poor districts has been countered by one backward step."³⁴ On the

26. *Id.*

27. Act of Mar. 9, 1966, ch. 24, pt. III, §§ 6–16, 1966 Ky. Acts 235, 235–37 repealed and reenacted by KERA, ch. 476, 1990 Ky. Acts 1208 (codified at KY. REV. STAT. ANN. §§ 160.605, 160.613, 160.621 (Michie/Bobbs-Merrill 1994)).

28. Act of Mar. 29, 1976, ch. 93, 1976 Ky. Acts 199, repealed by KERA, ch. 476, § 616, 1990 Ky. Acts 1208 (codified as amended in scattered sections of KY. REV. STAT. ANN., ch. 156–165 and other scattered chapters).

29. *Rose v. Council for Better Educ., Inc.*, 790 S.W.2d 186, 196 & n.10 (Ky. 1989).

30. Act of Feb. 13, 1979, ch. 25, § 6, 1979 Ky. Acts 33, 38–40 (codified as amended at KY. REV. STAT. ANN. § 160.470 (Michie/Bobbs-Merrill 1994)).

31. CAL. CONST. art. 13A. For a discussion of the anti-tax movement that spawned Proposition 13, see *California Referendum on Taxes Symptom of a Nationwide Revolt*, N.Y. TIMES, June 4, 1978, at A1.

32. See 1979 Ky. Acts 33, 38; *Council for Better Educ.*, 790 S.W.2d at 195.

33. Keating, *supra* note 10, at 8.

34. *Council for Better Educ.*, 790 S.W.2d at 196.

eve of *Council for Better Education* and KERA, property wealth per pupil varied from \$39,138 to \$341,707.³⁵ Levied equivalent tax rates ranged from 22.9¢ to 111.9¢.³⁶ Local revenue per pupil ran from \$80 to \$3716,³⁷ and average per-pupil expenditures for instruction varied from \$1499 to \$3709.³⁸

B. Early Quality and Accountability Measures

Kentucky also made some limited efforts to improve quality and build accountability into the public education system in the years before KERA. In 1978, the General Assembly created a statewide student assessment system.³⁹ Although this assessment system lacked the high academic standards later enacted by KERA, it did mandate basic skills in a number of core subjects for grades three, five, seven, and ten and required schools to submit annual school improvement plans to the state.⁴⁰ In 1984, the legislature enacted tougher legislation that required the State Board of Education to establish minimum standards for student performance and specified remedial steps to be taken by local school districts determined to be educationally deficient.⁴¹ These steps, however, produced only average performance by students. For example, on the 1990 National Assessment of Educational Progress (NAEP) eighth grade math assessment, Kentucky ranked approximately twenty-eighth out of participating states and territories.⁴²

As one case study of Kentucky education reform efforts concludes, "[t]his 'one step forward, one step back' approach to education reform suggests that social, political, and legal forces

35. Jacob E. Adams, Jr., *School Finance Reform and Systemic School Change: Reconstituting Kentucky's Public Schools*, 18 J. EDUC. FIN. 318, 331 (1993).

36. *Id.*

37. *Id.*

38. *Id.*

39. The Education Improvement Act of 1978, ch. 151, 1978 Ky. Acts 354 (codified as amended in scattered subsections of KY. REV. STAT. ANN. § 158 (Michie/Bobbs-Merrill 1992 & Supp. 1994)), *repealed in part* by KERA, ch. 476, 1990 Ky. Acts 1208.

40. *Id.*

41. Act of Apr. 9, 1984, ch. 347, 1984 Ky. Acts 805 (codified in scattered subsections of KY. REV. STAT. ANN. § 158 (Michie/Bobbs-Merrill 1992 & Supp. 1994)), *repealed in part* by KERA, ch. 476, 1990 Ky. Acts 1208.

42. See NATIONAL CTR. FOR EDUC. STATISTICS, U.S. DEPT OF EDUC., DIGEST OF EDUCATION STATISTICS 120 tbl. 117 (1993) (providing statistics for forty-one states, the District of Columbia, Guam, and the Virgin Islands).

were never truly in sync."⁴³ While many citizens of Kentucky wanted change, others seemed satisfied with the current system. Kentucky schools needed a galvanizing force to make the state recognize the moral imperative and practical necessity of improving its schools. The *Council for Better Education* decision would serve that role.

C. Modern Reform: The Council for Better Education Decision

Given the legislature's intermittent concern for the quality of the schools and creation of barriers against adequate funding, it is not surprising that Kentucky education reformers turned to the courts for assistance. The origin of *Rose v. Council for Better Education*⁴⁴ dates to 1984, when a number of local districts formed the Council for Better Education and, with encouragement and direction from Kentucky educator Arnold Guess and finance experts Kern Alexander and Richard Salmon, decided to pursue a lawsuit.⁴⁵ The group's efforts acquired significant credibility when former Governor and federal appellate court judge, Bert T. Combs, agreed to be the group's representative.⁴⁶ In 1985, the Council, representing sixty-six of Kentucky's 178 local school districts, filed suit against the current Governor, the Superintendent of Public Instruction, and the General Assembly.⁴⁷

Four years later, the Kentucky Supreme Court handed down its decision in favor of the plaintiffs.⁴⁸ The court held that the entire public school system was unconstitutional because the General Assembly had failed to maintain an "efficient system of common schools," as required by section 183 of the Kentucky Constitution.⁴⁹ Section 183 states that "[t]he General Assembly shall, by appropriate legislation, provide for an efficient system of common schools throughout the State."⁵⁰ The

43. Ronald G. Dove, Jr., *Acorns in a Mountain Pool: The Role of Litigation, Law and Lawyers in Kentucky Education Reform*, 17 J. EDUC. FIN. 83, 87 (1991).

44. 790 S.W.2d 186 (Ky. 1989).

45. See Dove, *supra* note 43, at 88-89.

46. *Id.* at 90.

47. *Id.* at 93-94.

48. *Council for Better Educ.*, 790 S.W.2d at 186.

49. *Id.* at 211-213.

50. KY. CONST. § 183.

court interpreted section 183 to establish three duties. First, the constitution assigns the duty of maintaining the school system solely to the General Assembly.⁵¹ Thus, the legislature must take whatever steps necessary to ensure the quality of the schools—it cannot simply pin the blame on local districts. Moreover, the legislation not only must establish the school system, “it must [also] monitor [the school system] on a continuing basis so that it will always be maintained in a constitutional manner.”⁵² Second, the system of common schools must be provided throughout the state.⁵³ Third, the system must be efficient.⁵⁴

The court proceeded to construct a detailed definition of an “efficient system of common schools,” by articulating nine requirements which incorporate section 183’s provisions.⁵⁵ In sum, the court declared that the General Assembly alone is obligated to establish, maintain, and fund a system of common schools free to all.⁵⁶ In addition, the court stated that common schools must be substantially uniform across the state and that all Kentucky children must have an equal opportunity to receive an adequate education, regardless of their place of residence or economic circumstances.⁵⁷ Furthermore, the General Assembly must provide sufficient funding to ensure that each child receives an adequate education.⁵⁸ Finally, and most significantly, the court defined an “adequate” education not in terms of resources, but as an education in which the child achieves seven “capacities”:

- (i) sufficient oral and written communication skills to enable students to function in a complex and rapidly changing civilization;
- (ii) sufficient knowledge of economic, social, and political systems to enable the student to make informed choices;
- (iii) sufficient understanding of governmental processes to enable the student to understand the issues that affect his or her community, state, and nation;
- (iv) sufficient self-knowledge and knowledge of his or her mental and

51. *Council for Better Educ.*, 790 S.W.2d at 203.

52. *Id.* at 211.

53. *Id.* at 208.

54. *Id.*

55. *Id.* at 212–13.

56. *Id.* at 212.

57. *Id.*

58. *Id.* at 213.

physical wellness; (v) sufficient grounding in the arts to enable each student to appreciate his or her cultural and historical heritage; (vi) sufficient training or preparation for advanced training in either academic or vocational fields so as to enable each child to choose and pursue life work intelligently; and (vii) sufficient levels of academic or vocational skills to enable public school students to compete favorably with their counterparts in surrounding states, in academics or in the job market.⁵⁹

The court concluded that the present system failed to meet these standards in three ways. First, the overall system was inadequate when compared to regional and national standards.⁶⁰ The court noted that Kentucky was near the bottom in every regional and national measure of educational resources and quality.⁶¹ Second, a great disparity in educational opportunities existed throughout the state with regard to curriculum, facilities, materials, teacher-student ratios, and other measures.⁶² Third, financial resources throughout the state were unequal and inadequate due to interdistrict variances in population and wealth.⁶³ The court stated that the state's two aid programs had not solved the problem: the MFP had provided aid solely on a per capita basis, achieving no equalization at all, and the PEP was not funded sufficiently to provide a significant amount of equalization.⁶⁴

The court suggested that, while the system must be adequately funded, the legislature was not required to raise taxes.⁶⁵ If property taxes were used to finance schools, however, property must be assessed at fair market value, and a uniform tax rate must be established statewide to ensure that all property owners were making a comparable effort.⁶⁶ The court also did not order compensatory aid for disadvantaged districts, as did

59. *Id.* at 212.

60. *Id.* at 197, 213.

61. *Id.*

62. *Id.* at 198, 213.

63. *Id.* at 198-99, 213.

64. *Id.* at 198-99. It is interesting to note that the court, while articulating a seven-part objective standard, also presented the evidence of inefficiency in comparative and equity terms. *See id.* at 196-99, 213. This may suggest that the court would have been satisfied if the state simply eliminated inequalities in the system.

65. *See id.* at 212, 216.

66. *Id.* at 216.

the New Jersey Supreme Court in a widely noted decision.⁶⁷ It did, however, require the legislature to fund the system sufficiently to ensure that every student receive an adequate education.⁶⁸ The court withheld the finality of the decision until ninety days after the adjournment of the General Assembly's 1990 regular session to allow the legislature a chance to respond.⁶⁹

Council for Better Education has been regarded as "one of the most dramatic and significant constitutional law decisions in the history of Kentucky."⁷⁰ The decision marked a major change in the course of education policy and finance in the state. The Kentucky court was the first in the nation to declare an entire state school system unconstitutional on the grounds that it failed to provide an *adequate* education to all elementary and secondary school students.⁷¹ The court also abandoned its traditional deference to the General Assembly in the area of education, and, despite its protests that it was merely interpreting constitutional text, it established a clear direction for the state's education policy.

The court broke new ground in education jurisprudence when it defined "adequate education" in terms of student "capacities."⁷² Perhaps most importantly, the court made the General Assembly directly accountable for ensuring that students acquire these capacities. It could no longer pass the responsibility or blame to the local school districts. This change made education a state-level issue and compelled the General Assembly to act. Yet, because the court articulated the outcomes so broadly and found educational inadequacy largely on the basis of interdistrict disparities in spending and resources, it remains

67. See *Abbott v. Burke*, 575 A.2d 359, 408 (N.J. 1990) (ordering that the law "be amended, or new legislation passed, so as to assure that poorer urban districts' educational funding is substantially equal to that of property-rich districts").

68. *Council for Better Educ.*, 790 S.W.2d at 212.

69. *Id.* at 216.

70. Bert T. Combs, *Creative Constitutional Law: The Kentucky School Reform Law*, 28 HARV. J. ON LEGIS. 367, 369 (1991).

71. State courts have generally declared school finance systems unconstitutional on equity grounds, i.e., that all children in the state should have equal access to educational resources through equalized spending or a guaranteed tax base. *E.g.*, *Serrano v. Priest*, 487 P.2d 1241 (Cal. 1971); *Helena Elementary Sch. Dist. No. 1 v. State*, 769 P.2d 684 (Mont. 1989); *Edgewood Indep. Sch. Dist. v. Kirby*, 777 S.W.2d 391 (Tex. 1989), *mandamus proceeding*, 804 S.W.2d 491 (Tex. 1991); see also William H. Clune, *The Shift from Equity to Adequacy in School Finance*, 8 EDUC. POL'Y 376, 377 (1994) (discussing cases holding school finance systems unconstitutional).

72. See *supra* note 59 and accompanying text.

an open question whether the court will hold the state to a specific level of performance if substantial equality is achieved.⁷³ Still, the court did not present any criteria, other than student performance, for evaluating whether the legislature has fulfilled its constitutional duty. It did not direct the legislature to adopt specific policy measures or funding plans. This suggests that the court was truly interested only in results.

D. The Legislature's Response: KERA

While the *Council for Better Education* decision was a bombshell when it was handed down in June 1989, many were pessimistic about how the General Assembly would respond. According to Bert Combs, lead counsel for the plaintiffs: "People close to state government thought that the odds were about even that the General Assembly would either ignore the Court mandate or would give it lip service and then drag its collective feet in mock deference."⁷⁴ Therefore, some may have been more than a little surprised when the legislature passed KERA, the most comprehensive statewide reform plan to date. Some have argued that the legislature responded so quickly and boldly because the *Council for Better Education* decision gave them both the political cover and an intellectual framework for improving the schools.⁷⁵ The decision also may have "imbue[d] the citizenry with a collective pride in ownership" and represented "an external force authorizing an important social change that the people intuitively knew was morally necessary and long overdue."⁷⁶ Others credit reform to the work of key politicians and various citizen groups like the Prichard Committee for Academic Excellence.⁷⁷ Another important factor may have been that neither the plaintiffs nor the court aimed to redistribute wealth directly from rich to poor districts,⁷⁸ although what the plaintiffs did seek, state equalization funded

73. See *supra* note 64.

74. Combs, *supra* note 70, at 375.

75. E.g., Kern Alexander, *The Common School Ideal and the Limits of Legislative Authority: The Kentucky Case*, 28 HARV. J. ON LEGIS. 341, 343 (1991) ("The court provided the legislature with both the nerve and the rationale to raise taxes, equalize school funding, and make other necessary changes.").

76. *Id.* at 344.

77. Dove, *supra* note 43, at 109.

78. *Id.* at 115.

by income taxes, redistributed funding indirectly. Instead, the court's mandate to the legislature was to improve quality across the board, using funding only as necessary to achieve improvement.⁷⁹ Finally, others see the reform as simply representative of the paradoxical nature of the state: "Kentucky often does the unusual."⁸⁰

In the spring of 1990, less than one year after *Council for Better Education* was decided, the General Assembly passed KERA.⁸¹ KERA completely abolished the state school system and created an entirely new one in its place.⁸² It contains three major reform components: curriculum,⁸³ governance,⁸⁴ and finance.⁸⁵ In the area of curriculum, KERA mandates a combination of goals, assessment, and accountability.⁸⁶

The legislation establishes six broad learning goals,⁸⁷ from which the state developed fifty-seven "academic expectations"⁸⁸ and a model curriculum framework.⁸⁹ KERA assumes that all students can learn⁹⁰ and expects all students to perform at a certain level.⁹¹ Thus the outcomes are intended to be challenging, "world class" standards,⁹² and not the basic skills that previous state tests have measured.

To measure progress toward the goals, KERA directs the state to develop a performance-based assessment for students in grades four, eight, and twelve.⁹³ The test was first administered in the 1991-1992 school year and results were reported in relation to the four performance standards established by the

79. See *Rose v. Council for Better Educ., Inc.*, 790 S.W.2d 186, 211-12 (Ky. 1989).

80. Combs, *supra* note 70, at 377.

81. KERA, ch. 476, 1990 Ky. Acts 1208 (codified as amended in scattered sections of KY. REV. STAT. ANN., chs. 156-165 and other scattered chapters).

82. See *id.*

83. *Id.* at pt. I.

84. *Id.* at pt. II.

85. *Id.* at pt. III.

86. *Id.* at pt. I.

87. KY. REV. STAT. ANN. § 158.6451 (Michie/Bobbs-Merrill 1992 & Supp. 1994). For a discussion of these six goals, see *infra* text accompanying note 132.

88. KENTUCKY DEPT OF EDUC., KIRIS ACCOUNTABILITY CYCLE I TECHNICAL MANUAL 6 (1995).

89. *Id.*

90. See, e.g., KY. REV. STAT. ANN. § 158.6455 (Michie/Bobbs-Merrill 1992 & Supp. 1994) ("It is the intent of the General Assembly that schools succeed with all students . . .").

91. See *id.* § 158.6453 (providing for an assessment "program to ensure school accountability for student achievement").

92. See KENTUCKY DEPT. OF EDUC., *supra* note 88, at 65 (standards were "first labeled 'world-class'" but "'world class' was discarded . . . because of . . . confusion").

93. KY. REV. STAT. ANN. § 158.6453.

State Board of Education: novice, apprentice, proficient, and distinguished.⁹⁴ During 1991–1992, approximately ten percent of the state's students tested at the proficient level in reading, math, science, and social studies.⁹⁵ Eventually, schools will be rewarded or sanctioned based on the results of the assessment.⁹⁶

KERA contains a number of other instructional reforms, including expanded pre-kindergarten and parent education programs, ungraded classrooms for the first four years of primary school,⁹⁷ and the introduction of technology on a state-wide level into the schools.⁹⁸ The state also has provided funds for extended day and summer programs for students seeking additional help in achieving the standards.⁹⁹

KERA significantly altered school governance as well. At the state level, the act transferred the duties of the elected State Superintendent of Public Instruction¹⁰⁰ to an appointed Commissioner of Education,¹⁰¹ and reorganized the State Department of Education.¹⁰² At the local level, each school established a school council with control over curricular and instructional matters.¹⁰³ Each council consists of three teachers, two parents, and the school's principal.¹⁰⁴

The legislature also enacted a number of measures intended to reduce political influences in staff hiring. Local school boards

94. KENTUCKY DEPT OF EDUC., *supra* note 88, at 65.

95. KY. OFFICE OF EDUC. ACCOUNTABILITY, *supra* note 3, at 65. "Proficient" means having achieved "solid competence" in a content area. KENTUCKY DEPT OF EDUC., *supra* note 88, at 65.

96. KY. REV. STAT. ANN. § 158.6455. These awards and sanctions create many difficult challenges for schools. For further discussion see *infra* Part II.E.

97. See KY. REV. STAT. ANN. § 158.030 (Michie/Bobbs-Merrill 1992) (defining primary school program); KY. OFFICE OF EDUC. ACCOUNTABILITY, *supra* note 3, at 148 (explaining restructuring of curriculum "to an outcomes based program" from kindergarten to third grade).

98. KY. REV. STAT. ANN. § 156.666 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

99. KY. REV. STAT. ANN. § 158.805 (Michie/Bobbs-Merrill 1992) (providing a school improvement fund); see also KY. OFFICE OF EDUC. ACCOUNTABILITY, *supra* note 3, at 90–92 (describing the Extended School Services Program which provides "additional instructional time and related services to needy students").

100. KY. REV. STAT. ANN. § 156.120 (Michie/Bobbs-Merrill 1992 & Supp. 1994) (establishing the duties of the superintendent).

101. KY. REV. STAT. ANN. § 156.147–.148 (Michie/Bobbs-Merrill 1992) (outlining the appointment and the duties of the commissioner).

102. KY. REV. STAT. ANN. § 156.016 (Michie/Bobbs-Merrill 1992).

103. KY. REV. STAT. ANN. § 160.345 (Michie/Bobbs-Merrill 1994) (vesting powers of school-based decision making in the school council).

104. *Id.*

now may hire only the superintendent,¹⁰⁵ who exercises all other hiring decisions.¹⁰⁶ School board members and superintendents may not have relatives who are employed by the district, and principals may not have relatives who work in the same school.¹⁰⁷

KERA introduced substantial new money into the school system with the Support Education Excellence in Kentucky (SEEK) formula. SEEK guarantees an annually adjusted minimum base of spending throughout the state,¹⁰⁸ while at the same time requiring districts to make a minimum equivalent tax effort of 30¢ per \$100 of assessed value.¹⁰⁹ In addition, districts may choose to generate additional revenue through two methods—Tier I and Tier II. Under Tier I, districts are allowed to levy additional taxes that produce revenues up to 15% of the minimum guaranteed base.¹¹⁰ Districts which participate in Tier I with property wealth below 150% of the statewide average per pupil assessment receive equalization funds.¹¹¹ Under Tier II, districts may levy additional taxes producing revenues equal to 30% of the sum of the minimum guaranteed base and Tier I.¹¹² No equalization is provided for Tier II revenue. Districts may not generate any additional revenues above this amount.¹¹³ The state also provides categorical funds for preschool programs, technology, and other purposes.¹¹⁴

Local revenues also have increased significantly.¹¹⁵ In fact, while the state's share of total revenues has declined slightly since KERA's enactment, the local share has increased somewhat.¹¹⁶ Total state and local support for education jumped by 38% from 1989–1990 to 1991–1992.¹¹⁷ Some property-poor dis-

105. KY. REV. STAT. ANN. § 160.350 (Michie/Bobbs-Merrill 1994) (appointment procedures for the superintendent).

106. KY. REV. STAT. ANN. § 160.380 (Michie/Bobbs-Merrill 1994).

107. KY. REV. STAT. ANN. §§ 160.180, .380 (Michie/Bobbs-Merrill 1994).

108. KY. REV. STAT. ANN. § 157.360 (Michie/Bobbs-Merrill 1992 & Supp. 1994). The base amount was \$2305 in 1990–1991 and \$2420 in 1991–1992. Adams, *supra* note 35, at 332.

109. KY. REV. STAT. ANN. § 160.470 (Michie/Bobbs-Merrill 1994).

110. KY. REV. STAT. ANN. § 157.440 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

111. *Id.*

112. *Id.*

113. *Id.*

114. KY. REV. STAT. ANN. § 157.3175 (Michie/Bobbs-Merrill 1992) (preschool education); KY. REV. STAT. ANN. § 157.670 (Michie/Bobbs-Merrill 1992) (technology trust fund).

115. See KY. OFFICE OF EDUC. ACCOUNTABILITY, *supra* note 3, at 13 (reporting a 35.3% increase in local revenues between 1989–1990 and 1991–1992).

116. See *id.*

117. See *id.*

tricts experienced a 39% increase in spending, while wealthier districts saw a rise of 18%.¹¹⁸ Overall, KERA introduced \$490 million in new funds into local school districts in 1990–1991, 77% of which were state funds.¹¹⁹

In addition, KERA appears to be making substantial steps toward funding equity. In its first year, per-pupil state and local revenue increased by 18.6%, while the range of per-pupil revenue among districts decreased by 45.9%.¹²⁰ Further, there is more disparity in the distribution of state aid, with poorer districts receiving more funding than wealthier districts.¹²¹ Similarly, the disparity in tax rates decreased by 46.3%.¹²² Thus, since the enactment of KERA, spending by poor and wealthy school districts has begun to converge at a higher level of expenditure.

II. KERA'S IMPLEMENTATION

The distinguishing feature of *Council for Better Education* is not only that it mandates that all students achieve an adequate level of skills, but also that it requires the performance level of the entire system to shift toward higher performance standards. Accomplishing this goal requires the state to confront three important questions. First, what constitutes high performance and how will the schools convert the broad "capacities" provided by the Supreme Court of Kentucky into the everyday curriculum? Second, how will the state know when it has achieved these goals? Third, how will the state ensure that every school achieves these goals, and what will the state do if they do not?

Kentucky's response to these questions may be measured by examining the development and implementation of the new statewide assessment system. KERA is a comprehensive package of accountability, governance, and instructional reform. The central mechanism driving the system, however, is assessment. Assessment establishes goals for the system and keeps it on track toward those goals. It communicates instructions to parties at the ground level and determines whether those

118. *Id.* at 42.

119. Adams, *supra* note 35, at 336.

120. *Id.* at 337–38.

121. *Id.* at 337.

122. *Id.* at 338.

instructions have been followed. Because it is so central, assessment provides the means for understanding the obstacles to be overcome and the decisions to be made in building a high performance educational system.

A. The Assessment Program

KERA created a statewide, performance-based assessment program.¹²³ Originally, KERA established that students in grades four, eight, and twelve were to be tested each year in each major content area: reading, mathematics, science, social studies, and writing.¹²⁴ Based on their performance, students are rated according to four levels of achievement: novice, apprentice, proficient, and distinguished.¹²⁵ Ultimately, assessment will consider a combination of portfolios, small-group exercises, and open-ended essay questions.¹²⁶

KERA requires the establishment of an interim assessment program similar to the National Assessment of Educational Progress (NAEP).¹²⁷ The interim assessment program must measure reading, math, science, social studies, and writing.¹²⁸ This program must also provide statewide and national comparative data.¹²⁹ The most critical component of the assessment program, however, is the requirement that the State Board for Elementary and Secondary Education establish a primarily performance-based program no later than the 1995–1996 school year and as early as the 1994–1995 school year.¹³⁰

123. KY. REV. STAT. ANN. § 158.6453 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

124. *Id.* After the 1993–1994 school year, part of the assessment was administered to students in the eleventh grade, *see* KY. REV. STAT. ANN. § 158.6453 note (Baldwin 1994), and to students in the fifth grade, *see* KENTUCKY DEP'T OF EDUC., KENTUCKY INSTRUCTIONAL RESULTS INFORMATION SYSTEM (KIRIS): 1995–96 DISTRICT ASSESSMENT COORDINATOR IMPLEMENTATION GUIDE 10 tbl. 2-A (1995).

125. *See supra* notes 94–96 and accompanying text.

126. *See* BETTY E. STEFFY, *THE KENTUCKY EDUCATION REFORM: LESSONS FOR AMERICA* 43–44 (1993).

127. *Compare* KY. REV. STAT. ANN. § 158.6453 (Michie/Bobbs-Merrill 1992 & Supp. 1994) with National Assessment of Educational Progress Improvement Act, 20 U.S.C. § 221 (1994). While the Kentucky statute would have permitted the use of the NAEP itself, financial and logistical considerations and federal legislation authorizing the NAEP precluded this option.

128. *See* KY. REV. STAT. ANN. § 158.6453 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

129. *Id.*

130. *Id.* *See generally* C. Scott Trimble, *Ensuring Educational Accountability, in HIGH STAKES PERFORMANCE ASSESSMENT: PERSPECTIVES ON KENTUCKY EDUCATIONAL*

B. Goals and Curriculum

The principal goal of the Kentucky assessment program is to measure whether students are receiving an adequate education as defined by the Supreme Court of Kentucky in *Council for Better Education*. In this case, the court articulated seven capacities that broadly describe what students must know and be able to do.¹³¹

The statute directed specific student outcomes by setting six statewide goals for student achievement:

- a) Schools shall expect a high level of achievement of all students.
- b) Schools shall develop their students' ability to:
 - 1) Use basic communication and mathematics skills for purposes and situations they will encounter throughout their lives;
 - 2) Apply core concepts and principles from mathematics, the sciences, the arts, the humanities, social studies, and practical living studies to situations they will encounter through their lives;
 - 3) Become a self-sufficient individual;
 - 4) Become responsible members of a family, work group or community, including demonstrating effectiveness in community service;
 - 5) Think and solve problems in school situations and in a variety of situations they will encounter in life; and
 - 6) Connect and integrate experiences and new knowledge from all subject matter fields with what they have previously learned and build on past learning experiences to acquire new information through various media sources.
- c) Schools shall increase their students' rate of school attendance.
- d) Schools shall reduce their students' dropout and retention rates.

REFORM 37 (Thomas R. Guskey ed., 1994) (discussing procedures used in establishing performance goals under KERA).

131. See *supra* text accompanying note 59.

- e) Schools shall reduce physical and mental health barriers to learning.
- f) Schools shall be measured on the proportion of students who make a successful transition to work, post-secondary education, and the military.¹³²

Along with developing specific goals from which success may be assessed, it also is important for the state to communicate the curriculum to teachers. This communication may be accomplished in the following ways:

- a. State curriculum frameworks;
- b. State textbook adoption through a State Textbook Commission;
- c. Expanded professional development opportunities for teachers; and
- d. Education Professional Standards Board to oversee teacher certification and training.¹³³

C. Measuring Student Achievement

The purpose of an assessment in a high-stakes accountability system,¹³⁴ however, is not merely to monitor school effectiveness; it is also a key component in affecting change. Because consequences for the district are tied to student performance on the assessment, instruction and learning will conform to the content and method of the assessment. Thus, the assessment is a powerful tool for raising student achievement through improving instruction. At the same time, any assessment program should be designed very carefully so that teachers employ the best instructional practices and students learn what they should learn.

1. *Kentucky's Reliance on Performance-Based Assessments*—To accomplish these goals, Kentucky has implemented

132. KY. REV. STAT. ANN. § 158.6451 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

133. See Adams, *supra* note 35, at 325–26 (proposing a strategy including uniform curriculum, professional development, assessment, and accountability).

134. KERA introduces high stakes for schools by providing for consequences to schools that are connected to student achievement. KY. REV. STAT. ANN. § 158.6455 (Michie/Bobbs-Merrill 1992 & Supp. 1994); see *infra* text accompanying notes 226–27.

performance-based assessments.¹³⁵ Performance-based assessments differ from traditional multiple-choice tests in that their items are designed to probe students' understanding more deeply and to resemble more closely exercises that students might perform everyday in the classroom. In mathematics, for example, instead of choosing among four or five possible answers, students solve the problems on paper and include a written explanation of their answers.¹³⁶ Performance assessments include writing and mathematics portfolios, which are collections of the students' best work.¹³⁷

The trade-offs of the two forms of assessment are quite straightforward and have been understood among teachers for many years. While a multiple-choice test can survey a student's level of specific knowledge, a single open-ended item in a performance-based assessment can measure a student's ability to organize knowledge on a particular matter and to communicate, usually in written form, a level of understanding.¹³⁸ The two forms of assessment may in fact require similar kinds of background knowledge, but the method of applying this knowledge may be very different. Teachers often acknowledge this difference when they construct their own classroom tests. Teacher-designed classroom tests may award, for example, one point of credit for a multiple-choice item and five to ten points for an essay question.

2. *Measuring Content Validity*—Performance assessments and traditional multiple-choice tests are equally useful in achieving content-related validity, i.e., in assuring the right skills are assessed.¹³⁹ Content validity typically is accomplished by using juries of content area experts to review the assessments for the purpose of certifying, for example, that a set of mathematical items measures the content intended to be measured.¹⁴⁰ The state used this process in developing the

135. See KY. REV. STAT. ANN. § 158.6453 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

136. See, e.g., KENTUCKY DEP'T OF EDUC., *supra* note 88, at 32–33 (requiring mathematics testing to measure ability to “solve problems by selecting and using appropriate strategies, models, and relevant mathematical content; and verify and generalize solutions by reflecting critically on the problem”).

137. See, e.g., *id.* at 41 (“Kentucky writing and mathematics portfolios are compilations of five to seven best pieces of student work . . .”).

138. See STEFFY, *supra* note 126, at 41–44 (explaining that the goal of KERA is to transform assessment from a “content-focused” system to one that demonstrates acquired “valued outcomes”).

139. KENTUCKY DEP'T OF EDUC., *supra* note 88, at 201. “Validity addresses the issue of whether a test measures what it is intended to measure . . .” *Id.* at 199.

140. *Id.* at 61.

Kentucky Instructional Results Information System (KIRIS) performance assessments.¹⁴¹ Committees of teachers from across the state were brought together for the purpose of advising the test development process.¹⁴²

Achieving content validity for the Kentucky assessments, however, will differ in an important way from the traditional task of achieving content validity, because Kentucky assessments are designed to raise learning above current levels. In traditional test development environments, the content validity review process would direct the content area experts to consider a survey of what is being taught within a discipline and to determine whether the proposed items actually measure the curriculum currently in place.¹⁴³ This process is well-defined because the function of assessment prior to the reform movement was to understand the curriculum as it was being offered and to measure that curriculum. In contrast, the aim of the new Kentucky assessments is to test students against "world-class" standards, rather than against the curriculum existing before the reform. Therefore, Kentucky's content experts must evaluate assessments against the newer and higher expectations.¹⁴⁴

3. *Measuring Consequential Validity*—The great advantage of performance-based assessments is that they can achieve the proper consequential validity.¹⁴⁵ They encourage instructional practices that will challenge students to apply a broad range of knowledge and skills to real and complex problems. Consequential validity will result because teachers prepare students to perform well on the tests used to evaluate themselves and their schools. Therefore, in planning assessment design, Kentucky's educators must consider the means by which they want teachers to teach students routinely. If the assessment employs open-ended essays, then teachers likely will use open-ended essays routinely.

Traditional multiple-choice assessments cannot encourage the kind of instructional practices KERA aims to implement, because they typically test only a sample of skills within a

141. *Id.* at 54.

142. *See, e.g.,* KY. OFFICE OF EDUC. ACCOUNTABILITY, *supra* note 3, at 69–73 (describing feedback from a rural Kentucky committee).

143. *See* KENTUCKY DEP'T OF EDUC., *supra* note 88, at 199.

144. *See id.* ("Traditional thinking about reliability and validity could lead to an assessment design contrary to the true purpose of KIRIS: supporting real change in the classroom.")

145. *See id.* at 215 (defining "consequential validity" as whether an assessment's "effects, or consequences, are aligned with its goals . . . [and lack] negative side effects").

curriculum and do not measure application in a realistic way.¹⁴⁶ Because tests are used for high-stakes accountability measures, such as public listings of school data, instructional programs must adjust to the specific tests used and, just as importantly, to the method of assessment.¹⁴⁷ When accountability measures are based solely on multiple-choice items, schools will devote resources, both financial and human, to instructional activities that teach students to view a problem as having several alternatives, one of which is a best response. If this instructional model were preferred, then the traditional modes of assessment and accountability probably were sufficient. While real statistical data is scarce, it was not uncommon to hear of instructional programs in which teachers focused their curriculum on assessment content, which, by design, was to have been a *sample* of the total curriculum.¹⁴⁸ Consequently, Kentucky's curriculum narrowed and its instructional program concentrated on strategies for taking multiple-choice tests.¹⁴⁹

In the same manner, Kentucky should encourage better instructional practices by basing assessment and accountability on performance-based components such as open-response items. For these items, students organize and produce written responses to particular problems, participate in group activities to gather data and understand problems before producing written responses, and produce relatively lengthy responses that are entered into a portfolio of best work.¹⁵⁰ Responses by teachers and students indicate that a positive change in instruction has been taking place since the adoption of the new

146. See STEFFY, *supra* note 126, at 47 ("There is growing, if far from universal, impatience with [standardized testing] that addresses chiefly facts and basic skills, leaving thoughtfulness, imagination, and pursuit untapped.") (citation omitted).

147. See KENTUCKY DEP'T OF EDUC., *supra* note 88, at 11 (describing the KIRIS accountability index and claiming that accountability data will "heighten public attention in order to focus schools on helping . . . students achieve . . . high standards").

148. See STEFFY, *supra* note 126, at 41 (depicting Kentucky's high school programs as "content-focused. . . with little linkage to the 'big ideas'").

149. See *id.* at 42. ("Districts relied on the State Department to dictate what tests would be used to assess student achievement, score the tests, and send the districts the results."). "[U]ntil quite recently, it was not common for a school district to . . . evaluate the effectiveness of the instructional program . . . other than through the results of . . . state achievement tests." *Id.*

150. See, e.g., KENTUCKY DEP'T OF EDUC., *supra* note 88, at 45 (providing an example of an open-response item where students are asked to create a graph and describe data before reaching a conclusion).

Kentucky assessments.¹⁵¹ In more and more classes, students must demonstrate their comprehension through open-response items, instead of simply recalling enough specific information to answer multiple-choice questions.¹⁵² From these developments, it is reasonable to predict that future instructional processes will focus less on helping students to select a correct response from a set of possible answers and more on discovering an answer and communicating it in a meaningful way.

4. *Accountability*—While some might argue that one system cannot achieve both instructional leadership and school accountability, the Kentucky model demonstrates that if both leadership and accountability are needed, they must be guided by the same assessment, or at least by assessments that are parallel in structure, content, and format.¹⁵³ Teachers cannot be asked to produce students who demonstrate, through written products and other modes of communication, their ability to apply knowledge to complex problems and, at the same time, be held accountable for student performance on traditional tests that do not require them to apply such knowledge. The kinds of routine instruction necessary to prepare students for success on traditional and performance-based assessments are not necessarily mutually exclusive, but rather are different in their emphasis and in the time devoted to different teaching strategies.

151. *E.g.*, STEFFY, *supra* note 126, at 52 ("Feedback from . . . students regarding . . . performance event assessment was quite positive. Statements such as, 'This is fun!' and 'Can't we do this more often?' were commonly reported."); KY. OFFICE OF EDUC. ACCOUNTABILITY, *supra* note 3, at 72-73 (reporting that one teacher appreciated the increased writing opportunities for students and another preferred performance assessment to traditional standardized testing).

152. KENTUCKY DEPT OF EDUC., *supra* note 88, at 64-65 (outlining the assessment program for the 1993-1994 school year and demonstrating reliance on open-ended questions, writing portfolios, and performance tasks). *But cf.* STEFFY, *supra* note 126, at 44.

At the present time, most secondary teachers in [Kentucky] have been unaffected by the reform. Generally, they use the same textbooks they used before, give the same type of tests, and grade students using the same point system. Few secondary teachers . . . have a thorough understanding of KERA . . . or the new state performance assessment system.

Id.

153. *See* STEFFY, *supra* note 126, at 47 (claiming that many state officials believe that in the ideal classroom, "it will be impossible to tell the difference between instruction and assessment").

5. *Teachers' and Parents' Anxiety Over Performance-Based Assessment*—When a new assessment that is oriented toward what should be taught and the way it should be taught is introduced, teachers undertake a new type of responsibility which may cause them significant anxiety.¹⁵⁴ Traditionally, accountability pressures have encouraged or forced educators to design their assessment to mirror the instruction they offer.¹⁵⁵ Such traditional assessment allows the instructors to know from past experience the content that they will be responsible for teaching.¹⁵⁶ Textbooks and content guidelines typically have defined this content.¹⁵⁷ When the assessment is designed to *lead* instruction, however, or build a test worth teaching to, in the initial phases, teachers will likely feel that they are not fully aware of how they should teach students in order to ensure achievement according to the new assessment design.¹⁵⁸ While Kentucky hopes that daily classroom experience and the content and format of assessment will become congruent over time, it is unlikely that this vision will be immediately obvious to classroom teachers or immediately implemented in the classroom.¹⁵⁹ The system should allow time for such consistency to evolve. Although teachers may not demand a traditional scope and sequence¹⁶⁰ of what is to be taught, they will demand an

154. See *id.* at 52–53 (“teachers are still struggling with identifying exactly what skills [should be] assessed” when evaluating students under performance-based assessment); KENTUCKY DEP’T OF EDUC., *supra* note 88, at 219 (admitting that “accountability assessment is stressful to many people”).

155. See STEFFY, *supra* note 126, at 41 (claiming that tests have “generally assess[ed] whether students . . . demonstrate knowledge of the content covered in the course”).

156. Edward H. Haertel, *Theoretical and Practical Implications*, in *HIGH STAKES PERFORMANCE ASSESSMENT: PERSPECTIVES ON KENTUCKY EDUCATION REFORM*, 65, 71 (Thomas R. Guskey ed., 1994).

157. STEFFY, *supra* note 126, at 41. (“The [Kentucky] curriculum has been largely textbook driven, dealing more with ‘covering’ the material than assuring student mastery of ‘big ideas’.”)

158. See KENTUCKY DEP’T OF EDUC., *supra* note 88, at 6–9 (providing educators with fifty-seven “academic expectations” so that teachers may “teach . . . toward the test” and the test will be worth “teaching to”).

159. See *id.* at 218 (recognizing that teachers will need additional training and professional development to prepare students for KIRIS assessments); see also STEFFY, *supra* note 126, at 44 (“Full integration of instruction with students learning at high levels and demonstrating competence through authentic performance may not happen within the professional lifetime of many . . . educators.”)

160. “Scope and sequence” is a term that describes support materials that traditionally came with textbooks or content guidelines that prescribed which skills to teach and in what order to teach them. In recent years, these were short multiple-choice or fill-in-the-blank tests. See STEFFY, *supra* note 126, at 42, 44, 59.

explanation of the assessment and a compatible curriculum designed to achieve the desired results.¹⁶¹

Performance-based assessments may produce anxiety among parents as well. Parents, at times, become concerned because they believe traditional multiple-choice assessments are more objective.¹⁶² Both parents and teachers may be concerned that the new assessment and the instructional reforms that accompany it may not prepare their children to do well on the traditional college entrance exams.¹⁶³ Studies show a positive correlation between performance on the American College Testing Program admissions test (ACT) and KIRIS, at least in the areas of reading, mathematics and science.¹⁶⁴ Still others are concerned about what values the teachers apply when scoring the open-ended responses and, therefore, what values will be promoted at school.¹⁶⁵ This concern has caused some parents to withhold their children from the new assessment.¹⁶⁶ This is not a trivial matter, because the Kentucky accountability model holds the school accountable for the performance of *all* students enrolled in public education. The state assigns students who have not taken the assessment to the novice level of performance for purposes of accountability.¹⁶⁷

161. See KY. OFFICE OF EDUC. ACCOUNTABILITY, *supra* note 3, at 69–72 (describing teacher dissatisfaction with inadequate information and training on KIRIS).

162. See Ben R. Oldham, *A School District's Perspective*, in HIGH STAKES PERFORMANCE ASSESSMENT: PERSPECTIVES ON KENTUCKY EDUCATIONAL REFORM 55, 57 (Thomas R. Guskey ed., 1994) (“[P]arents . . . seem unwilling to give up reliance on normative comparisons.”); see also STEFFY, *supra* note 126, at 59 (“Parents . . . want data that enables them to compare the achievement of their children with other children, and with children in other schools and districts.”).

163. See KY. OFFICE OF EDUC. ACCOUNTABILITY, *supra* note 3, at 71–72 (reporting that Orange County teachers believe that high school students need practice taking “nationally normed standardized tests to prepare them to do well on college entrance exams”).

164. KENTUCKY DEP'T OF EDUC., *supra* note 88, at 213 (describing a high correlation between the ACT and KIRIS in the 1991–1992 school year when KIRIS had a limited number of open-response items).

165. *Id.* at 219 (noting that the new assessment system “is controversial . . . because it makes clear assumptions and beliefs, many of which are different from those held by large segments of the population”).

166. Letter from C. Scott Trimble, Director, Kentucky Department of Education, Division of Accountability, to anonymous Kentucky school district superintendent 1 (Jan. 4, 1994) (on file with the *University of Michigan Journal of Law Reform*).

167. *Id.*

D. Components of the New Kentucky Assessments

As required by KERA, the KIRIS assessment system consists of two parts: a cognitive component and a noncognitive component.¹⁶⁸ The cognitive component measures progress toward Goals A and B of the statute and consists of the students' performance assessments.¹⁶⁹ The noncognitive component measures progress toward Goals C, D, E, and F of the statute.¹⁷⁰ Goals A and B will be accomplished throughout the total design of KIRIS because the standards for student performance are set high and the standards for schools focus on continued improvement.¹⁷¹

168. See KY. REV. STAT. ANN. § 158.6451 (Michie/Bobbs-Merrill 1992 & Supp. 1994) (enacting six goals for the Kentucky schools); see also Trimble, *supra* note 130, at 42-48 (categorizing these goals in terms of "cognitive and noncognitive indicators").

169. Goals A and B, respectively, are:

- (a) Schools shall expect a high level of achievement of all students.
- (b) Schools shall develop their students' ability to:
 1. Use basic communication and mathematics skills for purposes and situations they will encounter throughout their lives;
 2. Apply core concepts and principles from mathematics, the sciences, the arts, the humanities, social studies, and practical living studies to situations they will encounter throughout their lives;
 3. Become a self-sufficient individual;
 4. Become responsible members of a family, work group, or community, including demonstrating effectiveness in community service;
 5. Think and solve problems in school situations and in a variety of situations they will encounter in life; and
 6. Connect and integrate experiences and new knowledge from all subject matter fields with what they have previously learned and build on past learning experiences to acquire new information through various media sources.

KY. REV. STAT. ANN. § 158.6451.

170. Goals C, D, E, and F respectively, are:

- (c) Schools shall increase their students' rates of school attendance.
- (d) Schools shall reduce their students' dropout and retention rates.
- (e) Schools shall reduce physical and mental health barriers to learning.
- (f) Schools shall be measured on the proportion of students who make a successful transition to work, post-secondary education, and the military.

Id.

171. See *id.* (mandating that "[s]chools shall expect a high level of achievement for all students").

1. *Cognitive Indicators (Goals A and B)*—Because the new assessments represent a radical departure from Kentucky's past practices, a number of steps have been taken to minimize disruptions in the classroom and other technical problems that result from using cutting-edge assessment techniques.¹⁷² First, KERA allows a primarily performance-based assessment to be phased in over a five-year period.¹⁷³ Second, to meet certain assessment needs external to the Kentucky reform, the assessments have three methods of measurement. These three methods are: the KIRIS Transitional Assessment, the KIRIS Performance Events, and the KIRIS Portfolio Assessment.¹⁷⁴

Within the original design, the KIRIS Transitional Assessment was to be the primary assessment vehicle while the other methods would continue to expand and evolve.¹⁷⁵ Later, the Performance Events and Portfolio Assessments were to serve as the primary methods.¹⁷⁶ While each method of assessment will continue to be administered, the goal was that Portfolio Assessment eventually will be assigned the most weight in performance calculations because it can achieve the type of reforms in instruction and learning envisioned by KERA.¹⁷⁷ Based on experiences in the first years of implementation, the initial plans for assessment will likely undergo some fine tuning.

a. *The KIRIS Transitional Assessment*—The Transitional Assessment serves two major purposes. First, it is a test that may be quickly implemented to meet the immediate assessment needs of the school reform. Its use will afford educators the time to develop the technology and methodology needed for the performance-based assessments. Second, the Transitional Assessment will remain part of the state's assessment system

172. Trimble, *supra* note 130, at 44.

173. See KY. REV. STAT. ANN. § 158.6453 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

174. STEFFY, *supra* note 126, at 51–57 (describing the implementation of the three-part assessment program).

175. See *id.* at 52 (stating that the 1991–1992 Transitional Assessment test “was designed to establish a ‘baseline’ score for each school . . . and establish a target ‘threshold’ score for the 1993–94 assessment”); Trimble, *supra* note 130, at 44–45 (explaining that Transitional Assessment provides “the means for moving away from traditional multiple-choice . . . tests, toward [tests] based on . . . performance-based assessments”).

176. Trimble, *supra* note 130, at 46. This Article will refer to Performance Events and Portfolio Assessments collectively as “performance-based assessment.”

177. See *id.* (stating that “[t]he portfolio assessments represent the vision of KIRIS and the School Accountability Program”).

in order to meet other assessment needs that require more traditional methodologies, such as national comparative data.¹⁷⁸

The KIRIS Transitional Assessment contains both multiple-choice and open-ended items in the areas of reading, mathematics, science, and social studies. In each of these areas, the first test of the 1991–1992 school year consisted of forty common multiple-choice items administered to all fourth, eighth, and twelfth grade students, and a set of 180 matrix-sampled¹⁷⁹ multiple-choice items.¹⁸⁰ These items were administered in sets of fifteen and on twelve different forms of the assessment.¹⁸¹ In each of these content areas, there was a set of three common open-ended items administered on all tests and a set of twelve matrix-sampled open-ended items administered, one on each of the twelve forms of the assessment.¹⁸² The transitional writing component consisted of nine pairs of prompts administered in a matrix-sampled design from which the student was instructed to choose one of two prompts.¹⁸³ Each of the content areas was designed to be administered in ninety minutes, with students allowed an additional forty-five minutes to complete the assessment.¹⁸⁴ Thus, each section could have taken a total of 135 minutes to administer. The intent was to allow all students adequate time to complete the assessments.

Additionally, the Transitional Assessment is being implemented in order to meet some of the more traditional assess-

178. See KY. REV. STAT. ANN. § 158.6453 (Michie/Bobbs-Merrill 1992 & Supp. 1994) (mandating that the "interim testing program . . . shall be designed to provide the state with national comparisons").

179. "Matrix sampled" refers to a test design that calls for a certain percentage of the students tested to take different forms of the test. KENTUCKY DEP'T OF EDUC., *supra* note 88, at 52.

180. *Id.*

181. In the Kentucky program, each student takes a certain number of "common items" that are used to provide data on each individual student and statistically link performance on the different sets of matrix sampled items. *Id.* Each student also takes one of 12 possible forms, which means that each student takes only 1/12th of the total number of items in the pool of matrix-sampled items. *Id.* This method of sampling allows the expansion of the number of test items administered in a school far beyond what any one student could have taken and thus, greatly increases the breadth of content that can be assessed at any given time.

182. Trimble, *supra* note 130, at 44.

183. KENTUCKY DEP'T OF EDUC., *supra* note 88, at 37. During the first four years of the implementation, the writing prompts play no role in the accountability index calculations but have been used in writing portfolio analysis and reviews. See *id.* at 36. The writing portfolio has provided the writing data used in accountability calculations. *Id.*

184. *Id.* at 111.

ment needs of Chapter 1 programs.¹⁸⁵ Chapter 1 is a federally funded program which provides compensatory education services, typically designed to strengthen reading and mathematics programs, to schools that have a relatively large percentage of educationally deprived students.¹⁸⁶ While the school's eligibility for these federal funds is based on economic indicators, an individual student's eligibility for the services depends upon her educational need, demonstrated by either a standardized assessment or a systematic means of teacher review and identification.¹⁸⁷ Chapter 1 programs were required to measure their effectiveness by administering a standardized test that is uniformly applied to a particular grade level in a school district.¹⁸⁸ The KIRIS Transitional Assessment meets this requirement.¹⁸⁹

Kentucky's educators must consider both the goal of using the assessment to influence instructional practices and the confusion likely to result if the Transitional Assessment premises success on student performance on a multiple-choice assessment while most of the instructional program is held accountable for growth on a performance-based assessment. By retaining a limited multiple-choice segment, the Transitional Assessment hopefully will meet the requirements of the Chapter 1 program while also allowing the instructional program to focus on performance-based assessment. Still, there is a danger that teachers providing compensatory reading and mathematics services to students with demonstrated educational needs will spend too much time preparing these students to perform well on multiple-choice items and not enough time on teaching tasks that will increase performance on Kentucky's open-ended assessments. By focusing on multiple-choice items, teachers may avoid negative reviews by local educational agencies for failing to show growth within a Chapter 1 program.¹⁹⁰

185. "Chapter 1" refers to Title I of the Elementary and Secondary Education Act of 1965, Pub. L. No. 89-10, 79 Stat. 27 (codified as amended at 20 U.S.C. § 2701 (1988)). Congress has recently adopted amendments to Chapter 1 that took effect on July 1, 1995. See Improving America's Schools Act of 1994, Pub. L. No. 103-382, § 101, 108 Stat. 3519.

186. 20 U.S.C. § 2723 (1988) (defining eligible schools).

187. See *id.* § 2724 (determining eligible children).

188. *Id.*

189. See, e.g., KENTUCKY DEPT OF EDUC., COMPREHENSIVE TESTS OF BASIC SKILLS 1 (4th ed. 1989) (providing national comparison data for third grade Kentucky students) (on file with the *University of Michigan Journal of Law Reform*).

190. See 20 U.S.C. § 2731 (1988) (authorizing annual reviews of program effectiveness); see also COMMISSION ON CHAPTER 1, MAKING SCHOOLS WORK FOR CHILDREN IN POVERTY

Although both transitional and performance-based assessments deal with much of the same content, there is a difference in the type of instruction that each assessment encourages. For example, in the instruction of reading, past experience shows that teachers can boost student performance on multiple-choice tests by teaching students to respond to relatively short passages and to select the best answer from a list of possible correct responses. To do well on performance-based tests, however, students must become accustomed to reading lengthy selections and writing essays about them.¹⁹¹

Transitional Assessment will also serve as a link to the National Assessment of Educational Progress (NAEP) reading, mathematics, and writing assessments and thereby provide a source of national comparative data.¹⁹² This is a very important role, given that a major impetus to the *Council for Better Education* litigation and KERA was the conclusion that poor schools made Kentucky economically uncompetitive with the rest of the nation. For example, both the court in *Council for Better Education* and KERA set as a goal that Kentucky students "compete favorably with their counterparts in [other] states . . ." ¹⁹³ Through Transitional Assessment and NAEP, Kentucky will be able to compare its performance with the rest of the nation.

During the initial development of the Transitional Assessment, the relative roles of the multiple-choice and open-ended items were not clearly defined. The state decided to rely solely on the open-ended items and to ignore the multiple-choice items.¹⁹⁴ This decision was appropriate, given KERA's mandate to move toward a primarily performance-based assessment.¹⁹⁵

77-78 (1992) (proposing legislation that would authorize sanctions, including reductions in pay, dismissal, or transfer, for failure to show progress in a Chapter 1 program).

191. See Oldham, *supra* note 162, at 62. ("[Students] must be taught . . . how to develop assessment tasks [by] moving from using verbs such as 'list,' 'define,' and 'identify,' to verbs such as 'explain,' 'compare and contrast,' and 'defend.'").

192. The National Assessment of Educational Progress (NAEP) is the assessment program operated through contractors by the National Center for Educational Statistics. In the 1992-1993 school year, the Kentucky Department of Education requested that NAEP conduct a study in reading, mathematics, and writing. KENTUCKY DEP'T OF EDUC., *supra* note 88, at 211-12. The Department was then able to match NAEP scores in these content areas with their KIRIS Transitional Assessment. *Id.* at 212.

193. *Council for Better Educ.*, 790 S.W.2d 186, 212 (Ky. 1989); KY. REV. STAT. ANN. § 158.645 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

194. See KENTUCKY DEP'T OF EDUC., *supra* note 88, at 47 (stating that in 1994, the Kentucky Board of Education voted to eliminate multiple-choice questions from KIRIS, at least until the 1996-1997 school year).

195. See KY. REV. STAT. ANN. § 158.6453 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

Important to this decision was the goal that assessment reflect instructional practice more closely.¹⁹⁶

Prior to September 1994, the state had increasingly relied on the open-ended items and less on the multiple-choice items.¹⁹⁷ The multiple-choice items were relegated to a role of providing a safety net within the assessment model and data for use in Chapter 1 evaluation reports.¹⁹⁸ Within the matrix-sampled pool of open-ended items, interdisciplinary items were included.¹⁹⁹ By the 1993–1994 school year, the reading, math, science, and social studies content areas shared items with other areas such as the arts and humanities and practical living/vocational studies.²⁰⁰ Each student was tested on five common open-ended response items and two matrix-sampled items.²⁰¹ Twelve forms of matrix-sampled items are used.²⁰² Such testing results in a total of twenty-nine items in each content area except writing, where students continue to respond to one of a pair of writing prompts.²⁰³

b. KIRIS Performance Events Assessments—Performance Events were administered in math, science, and social studies during the 1991–1992 school year.²⁰⁴ During the 1993–1994 school year, certain of these events were made interdisciplinary, such as science and math, and math and social studies.²⁰⁵ In addition, during 1993–1994, a small number of arts and humanities, practical living, and vocational studies events were added.²⁰⁶ Performance events accomplish two important goals of the assessment: (1) causing instruction to make students apply skills to produce products that will be evaluated, and (2) emphasizing teams or groups of students working together.²⁰⁷

In performance events, students are asked to participate in certain exercises as a group and then to produce a written product based on that experience.²⁰⁸ For example, a group of four

196. See KENTUCKY DEP'T OF EDUC., *supra* note 88, at 47.

197. See *id.*

198. See *id.*

199. See *id.* at 42 (reporting that during the 1993–1994 academic year, “three tasks at each grade level were interdisciplinary”).

200. See *id.*

201. Trimble, *supra* note 130, at 45.

202. *Id.*

203. *Id.*

204. STEFFY, *supra* note 126, at 52.

205. See KENTUCKY DEP'T OF EDUC., *supra* note 88, at 73.

206. *Id.*

207. Trimble, *supra* note 130, at 45–46.

208. STEFFY, *supra* note 126, at 52.

students could be asked to observe and record data measuring how high balls made of different materials bounce when dropped from specified heights. Based on their observations, the group might be asked to produce certain data tables or other products.²⁰⁹ From this common experience, each student would provide responses whose accuracy would depend on how well the group had worked together to make observations and record data.

c. *KIRIS Portfolio Assessments*—The portfolio method of assessment represents the vision of KIRIS and its desired impact on instruction.²¹⁰ Ideally, assessment would not be separate from instruction: data would be gathered from within the instructional process using everyday student products from the classroom, rather than products designed only for assessment purposes.²¹¹ One may consider portfolio assessment to be a process of collecting a set of classroom artifacts and making some judgment about the quality of the instructional interaction between teacher and student that produced these artifacts.²¹² In the long run, portfolios may extend beyond items that can fit into a paper folder to such media as audiotape, videotape, and computer disk.

Performance assessment is in many ways similar to the method used by parents to monitor their child's progress. If school, district, and state data are to be compiled from this process, however, evidence of student achievement must be judged in a uniform manner against a common set of standards. Kentucky's implementation has focused on what has been referred to as a "best work" portfolio,²¹³ because the effectiveness of the school should reflect the teacher's ability to interact with a student over time, coaching the student to improve a particular piece of writing or a math project. This teacher-student interaction should better enable the student to apply those skills obtained from a particular discipline or a combination of disciplines. For example, a student will understand better why

209. Trimble, *supra* note 130, at 46.

210. KENTUCKY DEPT OF EDUC., *supra* note 88, at 151 ("In many respects portfolios are the portion of . . . KIRIS . . . that most directly and comprehensively supports educational reform . . .").

211. STEFFY, *supra* note 126, at 47.

212. KENTUCKY DEPT OF EDUC., *supra* note 88, at 151 ("Portfolios are closely integrated with . . . instruction . . . [P]ortfolios are developed by students over long periods of time—months and perhaps years—with considerable support and direction by teachers.").

213. *Id.*

and how to apply the skills learned in language mechanics instruction if the student applies punctuation rules to his own writing.

Portfolios present evidence from daily classroom activities of the student's current ability. Instruction and assessment should become one and the same, particularly from the student's perspective. In this way, Kentucky will (1) prevent the portfolio from becoming an added assessment burden placed on the classroom teacher;²¹⁴ (2) make teachers a meaningful part of school effectiveness; and (3) gather authentic achievement data about students.

Portfolios measure student ability more completely. Regarding past multiple-choice assessments, the Kentucky Department of Education (Department) has stated that such tests convey learning and achievement equivalent to one day of instruction and student interaction.²¹⁵ In contrast, teachers have an advantage of 174 additional instruction days, or rather, a working knowledge of student performance accumulated over the course of an academic year. The Department has urged parents not to analogize achievement with a traditional assessment without first consulting the teacher who has much more knowledge of student classroom performance.²¹⁶ The portfolio method, on the other hand, has the potential to capture some of this additional knowledge that only a teacher can have.

The 1991–1992 assessment included a writing portfolio consisting of five to seven “best work” pieces.²¹⁷ Students and teachers developed portfolio entries over the first six to seven months of the school year.²¹⁸ Through a statewide training program, the standards against which these portfolios were to be judged were communicated to classroom teachers, who then evaluated these portfolios.²¹⁹ During the 1991–1992 academic year, the teacher training was administered through a pyramid structure, with state and contractor trainers at the top of the pyramid and regional coordinators, local district cluster leaders,

214. *But cf.* Oldham, *supra* note 162, at 61 (warning that grading, and often regrading, portfolios will significantly add to teachers' workloads).

215. *See, e.g.*, KENTUCKY DEP'T OF EDUC., *supra* note 189, at 1 (advising that since scores are based solely on a single test given once a year, score reports should not be used as the only indicator of achievement).

216. *Id.*

217. *See* KENTUCKY DEP'T OF EDUC., *supra* note 88, at 41.

218. *Id.*

219. *Id.*

and classroom teachers forming the base.²²⁰ During the 1992–1993 academic year, the training model was modified to extend direct training from the state and contractor levels toward the base of the structure.²²¹ This was done to extend the consistency of training as far down the structure as possible. By the 1993–1994 school year, math portfolios had been added for grades four, eight, and twelve.²²²

The Kentucky accountability system must include all students, including those at risk of failure. Therefore, alternate portfolios were added during the 1992–1993 school year for the purpose of including these students in the assessment and accountability model.²²³ The alternate portfolios are designed to include those students who cannot meaningfully participate in the regular curriculum even with assistance, adaptive devices, and the instructional modifications available. At the high school level, these students are so severely handicapped that they are not expected to obtain a high school diploma.²²⁴ Because the accountability system strives to include all students in the assessment and accountability model, alternate portfolios are important, even though they are expected to affect no more than one percent of the student population.²²⁵

d. Evaluating Results—To construct a court-mandated, adequacy-based educational system like Kentucky's, the first challenge is to translate the broad goals established by the supreme court into specific student expectations and then to build an assessment and accountability model that both measures performance and encourages instructional reform. Once the model has been designed, KERA introduces high stakes for schools to provide the motivation necessary to meet the court's goals. These high stakes are consequences for school government that are connected to student achievement. Some of these high-stakes provisions will not go into effect until 1996.²²⁶ Schools that perform satisfactorily will be left alone; schools that do not may be supervised more closely by the state; schools that make significant improvement may receive financial rewards.²²⁷

220. *Id.* at 116.

221. *See id.*

222. *Id.*

223. *Id.* at 42.

224. *Id.* at 40.

225. *See id.*

226. *See id.* at 41 (basing sanctions on 1996–1998 biennial review).

227. KY. REV. STAT. ANN. § 158.6455 (Michie/Bobbs-Merrill 1992 & Supp. 1994) (providing for rewards and sanctions based on student improvement).

Because assessment results will carry such consequences, they probably will be heavily scrutinized by educators and the public, who will consider the results to be legitimate only if they accurately portray student achievement. This need for legitimacy raises a number of technical concerns, such as assessment reliability. Assessment reliability is a central technical issue in the administration of any testing program and a particularly high-visibility issue in an assessment program that has serious accountability implications. If schools, students, teachers, and administrators will be penalized or rewarded based on achievement results, then assessment must consistently yield an accurate measure of student performance, and more particularly, overall school performance.

The reliability of traditional assessments is measured by well-developed statistical methods,²²⁸ which are appropriate for calculating reliability statistics on relatively lengthy multiple-choice tests.²²⁹ Reliability statistics can range, in theory, from a value of zero, a perfectly unreliable assessment, to a value of one, a perfectly reliable assessment.²³⁰ Expectations for reliability statistics vary, depending on what is being measured, such as academic traits or attitudes. In traditional content areas, such as reading, math, science, and social studies, educators expect reliability statistics to be in the range of .90 and above.²³¹ Expectations vary with the grade level being assessed, with reliability expected to be greater in the upper grades and slightly lower in the primary grades.²³²

Performance-based assessment presents new challenges to measuring reliability, in that most performance-based assess-

228. See, e.g., JUM C. NUNNALLY, *PSYCHOMETRIC THEORY* 229-34 (2d ed. 1978).

229. "Lengthy" is a relative term. A traditional multiple-choice test may include 40 to 80 items that measure a content area such as reading. Lengthy tests obviously yield more data on which to base reliability calculations. Under its current design, KIRIS would yield 5 data points at the student level and 29 at the school level. See Trimble, *supra* note 130, at 45.

230. Reliability measures the likelihood of obtaining the exact same test score were it possible to test and retest a student under exactly the same conditions on multiple occasions. See NUNNALLY, *supra* note 228, at 191 ("Reliability concerns the extent to which measurements are repeatable In other words, measurements are intended to be stable over a variety of conditions in which essentially the same result should be obtained."). If administered numerous times to the same student, a test with a score range of 1 to 100 and an associated reliability of zero would yield a series of random scores. Although no perfect standardized student assessment exists, a test with a reliability of 1.0 administered to the same student numerous times would yield exactly the same score each time.

231. See *id.* at 226.

232. See KENTUCKY DEP'T OF EDUC., *supra* note 88, at 196.

ments will be based on relatively short tests.²³³ Performance assessments are at least as demanding cognitively and at least as long in duration as traditional multiple-choice assessments.²³⁴ While a student might respond to fifteen multiple-choice items in a fifteen minute period,²³⁵ the student may be able to respond to only one well-designed and engaging open-ended item in the same amount of time.

Performance assessments such as those designed and implemented in Kentucky during the 1991–1992 school year yielded reliability statistics for individual students that ranged from .58 to .79.²³⁶ Because additional items were added to the assessment in the 1992–1993 year, reliabilities increased to the .65 to .85 range.²³⁷ Recent analysis of the Kentucky data indicates that the administration of five-item, open-response tests can yield reliabilities approaching the .85 expectation.²³⁸

Reliability statistics are more than just numbers to enter into various technical documents; they indicate the degree of confusion that may occur in analyzing the data. Students who in fact have a strong command of the content being measured are very likely to score high on performance-based assessments to the same extent that they were likely to score high on traditional multiple-choice tests. In the first year of the assessment, the risk that a very bright student would score noticeably lower than expected was slightly higher with performance-based assessments than with traditional multiple-choice assessments. However, the probability of a student exceeding expectations was also higher. This perceived unreliability manifests itself when teachers and parents of students with high college entrance exam scores observe the child scoring poorly on KIRIS Transitional Assessments or when teachers observe students judged to be at the head of the class scoring poorly on

233. Reliability statistics partially depend on the number of items included in an assessment. All other components being held constant, the more items administered, the more reliable the test. NUNNALLY, *supra* note 228, at 243.

234. In most cases, performance assessments will actually demand more time and ability than the traditional multiple-choice assessment. See Trimble, *supra* note 130, at 45–46.

235. See KENTUCKY DEP'T OF EDUC., *supra* note 88, at 47.

236. *Id.* at 197 (providing reliabilities for open-response tests).

237. *Id.* During the 1993–1994 school year, reliabilities ranged from .66 to .85 for open-response tests. *Id.*

238. *Id.* (showing reliabilities of .83 and .85 in social studies at the eighth and twelfth grade levels and reliabilities above .75 in most other areas at all tested grade levels).

performance-based assessments. These cases are infrequent,²³⁹ but nonetheless, may be confusing to the teachers, parents, and students. When these apparent inconsistencies occur within a high-stakes accountability environment, the immediate reaction is often to challenge the credibility of the assessment device. Although complaints about reliability are not unique to performance-based assessments, the newness of such assessments may make them more suspect to some critics.²⁴⁰

If KERA based the consequences of performance on individual student performance, then performance-based assessment might be unreliable. The responsibility for performance on the assessment, however, is at the school level.²⁴¹ Within this larger unit, a greater number of data points can be used to measure reliability.²⁴² As a result, unreliable scores by individual students will tend to balance out: some will score too low, but others will score too high. For the 1992–1993 and 1993–1994 school years, composite or school reliabilities approached the .95 range.²⁴³ Therefore, the reliability at the school level, the level at which accountability decisions will be applied, is within the range of traditional expectations of reliability.

The performance events method may best illustrate the conflict between the desire to influence positively instructional practices in the classroom and the necessity of achieving statistically reliable indices. During the first two years of assessment, reliabilities for Performance Events seem almost random, ranging from 0.00 to 1.00 with a mean of approximately .30, an extremely low value from a traditional perspective.²⁴⁴ Focus groups of educators, however, have been reluctant to abandon this new form of assessment because of its positive affect in the classroom. It also may be that educators observe that Performance Events, which contain both a group problem-solving and

239. *But cf. id.* at 213 (revealing a high degree of correlation between the ACT and KIRIS but also admitting that in “a substantial number of cases” students score high on one and low on the other).

240. *See id.* at 219–20 (reporting that many Kentucky administrators, teachers, students, and parents do not believe that “all children can learn at high levels,” in accordance with KERA).

241. *See* KY. REV. STAT. ANN. § 158.6455 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

242. *See supra* note 233.

243. KENTUCKY DEPT OF EDUC., *supra* note 88, at 197. These reliabilities demonstrate an improvement from the 1991–1992 school year, in which reported school reliabilities approached the .90 range. *Id.* Using the most conservative assumptions, reliabilities resulted in values of approximately .85. More liberal assumptions result in values ranging between .90 and .95. *See id.*

244. *See* STEFFY, *supra* note 126, at 52.

individual response component,²⁴⁵ best emulate what many teachers consider to be progressive and sound instructional practices.²⁴⁶ Instructional leaders in local school districts may consider Performance Events as a significant contribution to the measurement methods that have been adapted from classroom instructional practices to meet measurement needs.

However, because Performance Events do not have the technical reliability exhibited by Transitional Assessment and because each student usually participates in only one of approximately sixteen events administered by a school, Performance Events do not produce individual student scores. Rather, they produce only school composite scores. During the first two years of the assessment, statistical reliability has not been a point of concern, but as accountability decisions become more imminent, questions of reliability may become more central.

One concern that has arisen regarding the Performance Events involves the question of who should administer the event. During both the 1992–1993 and 1993–1994 school years, Performance Events were administered by outside facilitators hired by the state.²⁴⁷ Outside facilitators were used rather than the local school staff to ensure security and to provide a degree of control over the assessment environment.²⁴⁸ Since results of the test carried significant consequences for the school, however, the school staff closely scrutinized the facilitators' procedures and lodged complaints about any irregularities that they thought would affect the test's reliability.

When teacher-scored portfolios are a major component of an assessment system, as they will be in Kentucky, questions regarding the ability of teachers across the state to score student products consistently against a common standard become very important.²⁴⁹ The state must provide a credible means of reviewing a subset of the portfolio scores and must prescribe actions to be taken if the state detects that some teachers are scoring inconsistently. A uniform set of student performance standards within any content area cannot be implemented and applied to teacher-scored portfolios quickly. Kentucky has

245. *See id.*

246. *See* Oldham, *supra* note 162, at 59 ("Many . . . educators in Kentucky expected the performance-based assessment . . . to be a driving force in the reform movement.")

247. *See* KENTUCKY DEPT OF EDUC., *supra* note 88, at 113–14.

248. *Id.*

249. *See id.* at 154–57 (describing the 1992–1993 "Inconsistent Scorer Study" which attempted to ensure consistency in scoring portfolios across Kentucky).

begun to identify those schools thought to have the greatest difficulty in applying the writing standards, to adjust those scores if in fact these difficulties were confirmed, and to provide the training necessary to assist those schools to understand and apply the standards better.²⁵⁰

The Kentucky review process included several phases intended to determine how consistently the writing standards were applied to the writing portfolios. First, the state identified certain portfolios from teachers whose scores were inconsistent with the rest of their district's scores.²⁵¹ These portfolios were rescored by select Kentucky writing teachers under controlled conditions.²⁵² Second, the state selected portfolios from a random sample of schools across the state for rescoring.²⁵³

The third phase occurred when a specific number of schools were selected for audit. During the 1992–1993 audit, locally assigned writing portfolio scores were compared against three other data sets: (1) on-demand writing tasks included in the KIRIS Transitional Assessment; (2) open-ended items in reading, math, science, and social studies on the KIRIS Transitional Assessment; and (3) the school's writing portfolio data from the previous year.²⁵⁴

This process did not, and should not, lead to the automatic conclusion that the school's portfolios were not scored to the specifications of the common standards, but rather that the scores differed enough from other data that a source external to the school should review the teacher-assigned portfolio scores to determine whether they were correct. If the scores are inconsistent with the statewide writing standards, then the audit procedure should attempt to make these scores consistent with the standards.²⁵⁵

Because all of the consequences of scoring the portfolios were not clearly understood throughout the state and because not all of the stakeholders found the audit procedure acceptable,²⁵⁶ these procedures continue to be reviewed and refined. The credibility of the audit process depends on more than the simple

250. See *id.* at 154–67 (detailing the mechanisms for rescoring inconsistent scores and training and monitoring inconsistent scorers to score with consistency).

251. *Id.* at 155.

252. *Id.* at 154–55.

253. *Id.* at 158–60.

254. *Id.* at 169.

255. See *id.* at 168 (stating that the purposes of the audit include bringing “identified schools into line as much as possible with Kentucky writing standards”).

256. *Id.* at 178.

scoring of portfolios by expert writing scorers—teachers must be trained adequately to score the portfolios. In addition, the audit process must address the standards' stability over time. As a school implements and refines a portfolio scoring training program, the audited teachers, in particular, may interpret clarifications in training procedures to mean that the portfolio scoring standards are shifting. While the experts may not view these clarifications as shifts in the scoring standards if they will result in more accurate scoring, such clarifications may in fact cause shifts in data. Probably the most critical component of making the audit credible to the schools is immediate and timely feedback to the scoring teachers affected by the audit. When the teachers of audited portfolios believe that they have scored the portfolios in good faith and to the best of their ability, these teachers probably will not accept adjustment of portfolio scores without a thorough explanation.²⁵⁷ In creating an audit procedure, the state did not intend merely to affect the method of scoring. Rather, the state intended that teachers would adapt teaching methods to scoring standards.²⁵⁸

e. Standards of Performance—A key feature of high-stakes accountability is that student scores on assessments are reported based on performance standards. On more traditional assessments, student achievement is based on normative scales such as percentile scores.²⁵⁹ These traditional scores communicate how one achieves compared to others and do not convey whether the performance is adequate, that is, whether students are able to perform as educated citizens.²⁶⁰ Perhaps this comparative means of reporting student achievement confuses the public's understanding of education, since at least fifty percent of the students should have scored at or above the fiftieth percentile, even though performance at or above the

257. *See id.* (explaining audit review procedures for schools that were "extremely critical" of the 1992–1993 audit results).

258. *See id.* at 168.

259. *See, e.g.,* EDWARD W. MINIUM ET AL., *STATISTICAL REASONING IN PSYCHOLOGY AND EDUCATION* 38–44 (3d ed. 1993) ("The percentile system is widely used in educational measurement to report the standing of an individual relative to the performance of a known group."). Percentiles are typically reported as national percentiles. *See id.* at 38–39. The test for which the score is reported is administered to a national sample of students in a particular year. *See id.* For example, a student scoring at the 53rd percentile in reading on a test standardized in the spring of 1990 performed better than 53% of the students who took the test in 1990. *See id.*

260. KENTUCKY DEP'T OF EDUC., *supra* note 88, at 6.

fiftieth percentile may not prove educational adequacy.²⁶¹ Many Kentuckians perceive that students are unable to comprehend written material or apply mathematical skills adequately to move them into technically demanding occupations.²⁶²

The intent behind establishing performance standards is to change the focus from measuring a student's achievement against other students to measuring how a student performs against an adequate level of performance.²⁶³ In Kentucky, four performance levels were established, "novice," "apprentice," "proficient," and "distinguished," with proficient deemed to be adequate performance.²⁶⁴ These names are not nearly as important as are their definitions and, more importantly, the agreement on what kinds of student products fulfill each of these standards. KERA expects high levels of performance from every student.²⁶⁵ As a result, "proficient" has been defined as a level of performance high enough to allow the student to be competitive in the economic and social environment of the twenty-first century.²⁶⁶

Because these standards are an integral part of a high-stakes accountability program, students and schools must see some tangible evidence when a student makes progress. And because the adequate level of performance represents a major increase above the current performance of students, an intermediate level between the novice standard and the proficient standard was needed. For this reason, the apprentice level was included.²⁶⁷ In addition, the "distinguished" standard was established to recognize the accomplishment of those who could exceed the proficient standard with better organized and more detailed work.²⁶⁸

2. *Noncognitive Indicators (GOALS C, D, and F)*—KERA established goals for Kentucky's school system in areas other than student performance. Accordingly, schools will be judged

261. As the standardization date becomes older and the state has administered the test several times, the actual number of students scoring above the 50th percentile is likely to increase because, among other reasons, instruction tends to drift toward the specific content of the test. See STEFFY, *supra* note 126, at 41-42.

262. See, e.g., KENTUCKY DEPT OF EDUC., *supra* note 88, at 4 (reporting that Kentucky residents support instruction that teaches students how to better handle real world experiences).

263. See *id.* at 6.

264. *Id.* at 65.

265. KY. REV. STAT. ANN. § 158.6451 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

266. KENTUCKY DEPT OF EDUC., *supra* note 88, at 65.

267. See *id.*

268. *Id.*

on their ability to increase attendance rates,²⁶⁹ decrease dropout rates,²⁷⁰ decrease retention rates,²⁷¹ and increase the proportion of students who make a successful transition to adult life.²⁷² Although contributing approximately sixteen percent to the accountability index calculations,²⁷³ these data have received serious scrutiny by both local and state interest groups.²⁷⁴ Because these indicators vary less from school to school than do the cognitive indicators, one could argue that the real impact of these data is something less than sixteen percent. Nevertheless, including these factors in the accountability index arguably affected both definition of the index and data collection procedures.

Percent-in-attendance data has remained constant in both data definition and collection.²⁷⁵ This consistency was possible in part because attendance data were an important factor in past school funding procedures²⁷⁶ and therefore had been scrutinized carefully in the past for appropriate definition and data collection procedures.

Other indicators, however, have raised a number of issues. For accountability purposes, the scope of retention rates—the percentage of students that have not been promoted to the next grade level—has been restricted to grades four through twelve²⁷⁷ because KERA replaces what has been traditionally regarded as kindergarten, first, second, and third with an ungraded primary program within which “failure” is not supposed to occur.²⁷⁸ In other words, students should remain in the primary program until they are ready to enter the fourth grade. As opposed to past data collection procedures, schools now have the chance to consider the impact of summer school programs on the retention rate.²⁷⁹ Therefore, it is now necessary to collect this data in the fall following the current school year.²⁸⁰

269. KY. REV. STAT. ANN. § 158.6451(a).

270. *Id.* § 158.6451(b).

271. *Id.* § 158.6451(c).

272. *Id.* § 158.6451(f).

273. See KENTUCKY DEP'T OF EDUC., *supra* note 88, at 118.

274. Trimble, *supra* note 130, at 48.

275. *Id.*

276. *Id.*

277. *Id.*

278. See KY. OFFICE OF EDUC. ACCOUNTABILITY, *supra* note 3, at 148.

279. Trimble, *supra* note 130, at 48.

280. For example, students participating in summer school programs might be designated as “retained” in June but may qualify for promotion by the beginning of the following September. *Id.* at 48.

In addition, the Kentucky Department of Education has modified its definition of "dropout" to be consistent with the definition of the National Center for Education Statistics (NCES).²⁸¹ This definition includes summer dropouts and unverified transfers to other schools.²⁸² Schools also have an opportunity to identify students who seem to have dropped out and have since returned to school. Such students would not be counted a dropout if they had returned to an academic setting.²⁸³ In such circumstances, the dropout statistic will be modified to reflect these changes.²⁸⁴

Data on the transition to adult life²⁸⁵ continues to be consistent with Kentucky's historical definitions, but the method of collecting such data have been improved significantly.²⁸⁶ In essence, NCES's verification procedures for determining the whereabouts of potential dropouts have been adapted to determine the whereabouts of graduates.²⁸⁷ While there was a need to extend the successful transition to adult life over the longest period possible, the time frame had to be restricted to September 1 through November 15 in order to inform schools of their accountability index score in a timely fashion.²⁸⁸ If a graduate had entered the work force, the military, or higher education by the end of this period, then the graduate had made a successful transition to adult life.²⁸⁹

E. Constructing the Accountability Index

The bottom line for school improvement is the accountability index, which combines data from indicators from each of KERA's goals and weighs them to produce one number mea-

281. *Id.* NCES is part of the Office of Educational Research and Improvement of the U.S. Department of Education. 20 U.S.C. § 1221e-1 (1988).

282. Trimble, *supra* note 130, at 48.

283. *Id.* at 48-49.

284. *See id.* at 49.

285. *See supra* note 272.

286. Trimble, *supra* note 130, at 49.

287. *Id.*

288. *See id.*

289. KY. REV. STAT. ANN. § 158.6541 (Michie/Bobbs-Merrill 1992 & Supp. 1994); *cf.* KENTUCKY DEP'T OF EDUC., *supra* note 88, at 118 (defining "successful transition to adult life" as including at least 30 hours per week of full-time enrollment in post-secondary school, employment in a non-temporary position, active military status, or a combination of the three).

asuring the school's overall performance. Based upon this number the state figures its rewards or penalties for performance.²⁹⁰

The accountability index is the statistical "average of the cognitive and noncognitive indices for a school or school district."²⁹¹ A school's accountability index is comprised of a "baseline" score, or the school or school district's percentage of successful students at the beginning of each biennium, and an "improvement goal," set by the State Board for Elementary and Secondary Education (Board), which the school is expected to meet by the end of the following year.²⁹²

The index is calculated in the following way. The open-ended items of the KIRIS Transitional Assessment yield two indicators for each school of the percentage of students scoring at each of the four ability levels; novice, apprentice, proficient, and distinguished for the content areas of reading, math, science, and social studies.²⁹³ One measure is based on the commonly administered items in each content area, while the second is based on the matrix-sampled open-ended items.²⁹⁴ The administration of the performance events yields estimates of percentages of students scoring at each of the four standards in reading, math, science, and social studies.²⁹⁵ The writing portfolio yields this measure for the writing content area.²⁹⁶ These data are combined with percentages of attendance, retention rates, dropout rates, and the percentage of graduates making a successful transition to adult life.²⁹⁷

The Board had to select a means of weighing the different performance levels in the cognitive component. The Board chose a rather simple and straightforward method. The Board assigned a relative value to each performance standard by which the percentage of students scoring within each standard would

290. See KENTUCKY DEP'T OF EDUC., *supra* note 88, at 144.

291. 703 KY. ADMIN. REGS. 3:060 (1993).

292. See KENTUCKY DEP'T OF EDUC., *supra* note 88, at 11; see also KY. REV. STAT. ANN. § 158.6455 (Michie/Bobbs-Merrill 1992 & Supp. 1994) (defining "threshold" as the percentage of growth required for a school or school district by the State Board).

293. See HIGH STAKES PERFORMANCE ASSESSMENT: PERSPECTIVES ON KENTUCKY'S EDUCATIONAL REFORM app. at 91 (Thomas R. Guskey ed., 1994).

294. *Id.*

295. See *id.* app. at 91-92.

296. See *id.*

297. *Id.* app. at 92-93. Retention data and dropout data are inverted in the calculation to represent these factors in a positive direction. That is, they represent the percentage of students promoted and the percentage of students remaining in school. *Id.* app. at 93.

be multiplied.²⁹⁸ The first values proposed were 0 for Novice, 2 for Apprentice, 5 for Proficient, and 6 for Distinguished.²⁹⁹ The Board concluded that this method of weighing did not adequately reflect the value of performance at the Distinguished level and accordingly raised the relative value for the Distinguished to 7.³⁰⁰

Because of the appeal of scaling the accountability index so that its goal is a score of 100, a point at which all students might be thought of as successful, the Proficient level was ultimately assigned a 1.00 value.³⁰¹ This made the scale easier to interpret and more in line with the percentage scale suggested by KERA.³⁰² With this in mind, the Novice level was assigned a value of zero which established a clearly definable minimum on the accountability index scale. This reassignment resulted in the conversion of the above relative weights to those reported in Table 1 as they are used today to calculate the accountability index. To derive the contribution of each category to the index, these weights are multiplied by the percentages of students scoring within each standard.³⁰³ Table 2 summarizes the contribution of each method to the five content areas: reading, math, science, social studies, and writing. These contributions apply to grades four, eight, and twelve. Because the application of the four noncognitive factors differs for each of the grades, their contributions are summarized in Table 3.

The noncognitive indicators included in the index reflect a second set of complex value judgments. While the successful transition to adult life is valued highly by most educators, this factor is not a meaningful measure of elementary or middle school performance.³⁰⁴ Furthermore, most educators would probably agree that success in the adult world cannot be judged at a single point in time. To include the indicator in the accountability index, however, the data must be summarized at a particular point in time and be related to a particular class of seniors. Otherwise, the accountability process would have to be postponed until there had been adequate time to observe the

298. See *id.* app. at 92.

299. See *id.*

300. See *id.* at 50.

301. See KENTUCKY DEP'T OF EDUC., *supra* note 88, at 146.

302. See KY. REV. STAT. ANN. § 158.6455 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

303. HIGH STAKES PERFORMANCE ASSESSMENT: PERSPECTIVES ON KENTUCKY'S EDUCATIONAL REFORM, *supra* note 293, app. at 92.

304. *Id.* at 50.

success of students in adult life. Additional options, such as “lagging”³⁰⁵ this indicator one or more years, have been suggested. In addition, the data upon which the transition to adult life is based are not easily reviewed by sources external to the reporting schools and districts.³⁰⁶ For a combination of these reasons, this indicator contributes only 37.5% to the non-cognitive index for high school seniors.³⁰⁷

Similar concerns have surrounded the weighing of the dropout data. Dropout data cannot be defined meaningfully below the seventh grade level.³⁰⁸ Dropout prevention is a problem of greater importance at the high school level than at the middle school level, where retention rates are more important.³⁰⁹

TABLE 1
WEIGHTS ASSIGNED TO
PERFORMANCE STANDARDS*

| Performance Standard | Relative Values |
|---|-----------------|
| Novice | 0.0 |
| Apprentice | 0.4 |
| Proficient | 1.0 |
| Distinguished | 1.4 |
| *Trimble, <i>supra</i> note 130, at 50. | |

305. The strategy of lagging a particular noncognitive indicator would allow the data included in the 1993–1994 index calculations to be taken from the 1992–1993 school year. This method would allow for the timely calculation of accountability indices while providing more time to collect data concerning the successful transition to adult life.

306. According to Ben Oldham, former Director of Research and Assessment for Fayette County Public Schools, while verification of postgraduation status may come from “a variety of sources . . . most of the time, . . . data are collected through telephone calls to the student or the student’s parents.” Oldham, *supra* note 162, at 61.

307. 703 KY. ADMIN. REGS. 4:010 (1994).

308. Trimble, *supra* note 130, at 50.

309. *Id.* at 51.

TABLE 2
SUMMARY OF CONTRIBUTIONS OF EACH
CONTENT AREA (GRADES 4, 8, AND 12)*

| | Reading Index | Math Index | Science Index | Social Studies Index | Writing Index |
|-------------------------|---------------|------------|---------------|----------------------|---------------|
| Transitional Assessment | | | | | |
| Common Matrix | 50 | 40 | 40 | 40 | 0 |
| Performance | | | | | |
| Events Portfolio | 0 | 20 | 20 | 20 | 0 |
| | 0 | 0 | 0 | 0 | 100 |

*Trimble, *supra* note 130, at 51.

TABLE 3
SUMMARY OF CONTRIBUTIONS ON COGNITIVE
FACTORS TO THE NONCOGNITIVE INDEX
FOR GRADES 4, 8, AND 12 (PERCENTAGES)*

| | Grade 4 | Grade 8 | Grade 12 |
|--------------------------|---------|---------|----------|
| Attendance Rates | 80.0 | 40.0 | 20.0 |
| Retention Rates | 20.0 | 40.0 | 5.0 |
| Dropout Rates | 0.0 | 20.0 | 37.5 |
| Transition to Adult Life | 0.0 | 0.0 | 37.5 |

*Trimble, *supra* note 130, at 51.

There is no clear consensus on how retention should be considered, although it does appear to be a problem addressed at the middle school level.³¹⁰ Only attendance and retention data may be considered for the fourth grade,³¹¹ with attendance being the more important of the two, at least in the sense of fostering accountability. The nongraded primary program which replaces the traditional kindergarten, first, second, and third grades makes retention data below the fourth grade irrelevant.³¹²

Due to concerns surrounding the accuracy of this noncognitive data, the procedures used to collect such data during the 1991–1992 and 1992–1993 school years were studied by the Department of Education.³¹³ While the study found the collection procedures to be “thorough and reliable,” the Department decided to expand the number of options for districts to submit data, to standardize collection procedures and timelines, and eliminate duplication.³¹⁴

The calculations summarized here provide an index on which 100 represents a school that has brought all of its students to a proficient level of performance and has achieved perfection in all of the noncognitive factors: a 100% attendance rate, a 0% dropout rate, a 0% retention rate, and a 100% rate of successful transitions to adult life.³¹⁵ The five cognitive indices and the one noncognitive index were averaged to establish the 1991–1992 baseline accountability index, from which growth would be measured during the 1992–1994 biennium.

Success on the Kentucky index is not intended to represent an obtainable goal under current instructional delivery systems, but rather a future goal of an improved school system. The Kentucky Department of Education believes that a period of approximately twenty years is needed before adequate performance will be achieved.³¹⁶ This time frame has been estimated without knowing how quickly Kentucky's schools can adjust to this new kind of assessment. Several years of growth data will need to be analyzed before the state will know whether the twenty-year goal is reasonable.

310. *Id.*

311. *Id.*

312. See *supra* notes 277–78 and accompanying text.

313. KENTUCKY DEPT OF EDUC., *supra* note 88, at 118.

314. *Id.*

315. Trimble, *supra* note 130, at 51–52.

316. KENTUCKY DEPT OF EDUC., *supra* note 88, at 11.

1. *A Barriers To Learning Index*—One of the goals of KERA is the reduction of physical and mental barriers to learning.³¹⁷ Incorporating this goal into the accountability index has proven difficult. This problem has been considered mostly within the KIRIS Noncognitive Indicators Advisory Committee (NIAC).³¹⁸ The committee found it fairly easy to reach consensus as to the causes of mental and physical barriers to learning.³¹⁹ Such barriers include teenage pregnancy, drug abuse, alcohol abuse, child abuse, juvenile delinquency, poverty, single parent homes, transient or migrant families, and mental or physical disabilities.

While the public school representatives did not take issue with the importance of addressing the needs of these children, they questioned whether developing an index measuring the reduction of physical and mental barriers to learning would be the most effective means of assuring that these needs were addressed. In addition, those affiliated with the public schools tend to question what the public schools really can do to eliminate directly the causes of these barriers.

The public school representatives, however, based their resistance to a barriers index not on any one factor but rather on a complex notion of how such an index would affect the accountability index and how the special needs of certain students should best be met. In the forefront of this discussion was the problem of determining which data would be included in such an index, with what reliability, and at what cost to the school. Representatives of local education agencies are quite open to adding "defensible" data to an accountability index but also maintain that the value of such data must justify the additional burden that it will place on the school system. Value must be measured both in terms of the quality of data and in terms of the impact on the services delivered to the target population. If the creation and implementation of a barriers index draws resources away from the delivery of services, or has such little impact on the accountability index that there is no real impact on programs offered, then other alternatives should be con-

317. KY. REV. STAT. ANN. § 158.6451 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

318. See KENTUCKY DEP'T OF EDUC., *supra* note 88, at 11.

319. Mr. Trimble was the Kentucky Department of Education's staff to the KIRIS Noncognitive Indicators Advisory Committee (NIAC). References in the text to the NIAC and the information it learned are based on his knowledge and experiences with the committee.

sidered. In this case, resources diverted to such an index would not reflect the best use of the limited resources available to education.

The questions related to the data's value, quality, and cost arise from the difficulties of gathering meaningful data on barriers to learning, particularly at the school level. For example, the Kentucky Department of Education's ability to judge accurately the degree of drug or alcohol abuse within the school age population, as compared to the larger community, is suspect. While the barriers addressed may not be the same in every school and within every community, there should be some agreement as to which problems are of most concern throughout the state.

In addition, some local school district representatives believe that many existing programs and a great deal of effort are currently being directed at overcoming these barriers. A few examples are programs specially designed to meet the needs of teenage parents, alternative schools for students with discipline problems that prevent them from participating in a regular instructional setting, and dropout prevention programs.³²⁰ School officials might prefer to direct additional personnel and financial resources, if available, toward expanding and improving existing efforts to address these problems instead of toward collecting data for an accountability index.

Advocates for students faced with barriers to learning note that these programs merely react to existing problems and are not proactive.³²¹ But school staff are hard pressed to attack proactively the causes of such barriers as poverty and single parent homes. On the other hand, school districts may take some preventive steps against drug abuse and teenage pregnancy, for example, through alternative school curriculums.³²²

In addition, other components of the accountability index may serve as reasonable proxies for a direct measure of barriers to learning. If barriers to learning were being reduced, at least five things would be observed in the accountability index. First, achievement measured by the cognitive component of the assessment should increase. Second, a larger percentage of the total student population should achieve at high levels. Third,

320. See, e.g., 703 KY. ADMIN. REGS. 4:080 (1994) (designating alternative schools for at-risk students, such as drug abusers, physically abused students, discipline problem students, and "nontraditional" students, as "A5" schools).

321. See STEFFY, *supra* note 126, at 62-63.

322. See 703 KY. ADMIN. REGS. 4:080 (1994).

attendance rates should increase and retention rates should decrease. Fourth, dropout rates should decline. Fifth, successful transitions to adult life should increase. While a final decision is still under consideration,³²³ using the current accountability index as a proxy has been recommended to the Department. Under this option, the school improvement plan would include a mandatory review of barriers to learning and would base appropriate strategies to address these barriers on these reviews.

2. *Establishing Overall Performance Requirements for Schools*—Under KERA, schools will be rewarded or penalized based on the relative progress they make toward the established level of adequate performance, rather than on the actual achievement of adequacy.³²⁴ Progress will be measured in the following manner. A school's accountability index for Year One will serve as the baseline.³²⁵ A growth index will then be calculated by comparing the indices for Year Two against this baseline.³²⁶ The improvement goal, or the amount of growth a school will be expected to make by the end of each two-year accountability cycle, will be the difference between the school's baseline index and 100, with this difference divided by ten.³²⁷ This calculation is based on the assumption that it will take twenty years for schools to achieve the goals established by the State Board of Education.³²⁸

The Board had difficulty establishing the amount of growth required of schools for two reasons. First, the kinds of assessments used in Kentucky have been tried in various places but never in their current forms or under the pressures of high-stakes accountability.³²⁹ Because there is no relevant past experience against which to measure predictions, setting reasonable goals for growth is difficult.³³⁰ The Board felt that the threshold should establish the expectation that, in time, all students and all schools should be successful. The threshold was set such that were all Kentucky schools able to maintain

323. KENTUCKY DEP'T OF EDUC., *supra* note 88, at 11.

324. *See supra* notes 226–27 and accompanying text.

325. HIGH STAKES PERFORMANCE ASSESSMENT: PERSPECTIVES ON KENTUCKY'S EDUCATIONAL REFORM, *supra* note 293, app. at 93.

326. *Id.*

327. *Id.*

328. KENTUCKY DEP'T OF EDUC., *supra* note 88, at 11.

329. *See STEFFY, supra* note 126, at 91, 257 (noting Fort Worth's and Chicago's systems).

330. *See* KENTUCKY DEP'T OF EDUC., *supra* note 88, at 11.

the prescribed level of growth, all schools would obtain an accountability index of 100 within the next twenty years.³³¹ This expectation must be closely reviewed at the end of the next accountability cycle.

3. *The High Stakes: Consequences for Performance*—KERA states that “[i]t is the intent of the General Assembly that schools succeed with all students and receive the appropriate consequences in proportion to that success.”³³² This mandate sets the tone for the accountability program. KERA requires the Kentucky State Board of Education to establish a threshold that each school must meet by the end of the next biennium.³³³ Schools will receive financial rewards on behalf of instructional staff when the school exceeds its assigned threshold by at least one percent.³³⁴ If the school reaches but does not exceed its threshold by at least one percent, the school will be considered successful and will not be sanctioned.³³⁵ A school that maintains its current accountability index but fails to meet its threshold will be required to produce a school improvement plan describing how it will meet the threshold during the next biennium.³³⁶ Such a school is eligible for school improvement funds allocated by the General Assembly.³³⁷ A school in which the proportion of successful students declines by less than five percent is also eligible for school improvement funds and will receive the services of at least one distinguished educator to assist with planning requirements and implementation.³³⁸ Such a school will be considered a “school in decline.”³³⁹ Finally, if the proportion of successful students declines by five percent, then the school will be considered a “school in crisis.”³⁴⁰ For a school in crisis, all certified staff are placed on probation, parents are

331. *See id.* at 11.

332. KY. REV. STAT. ANN. § 158.6455 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

333. *See supra* Part II.E.2.

334. KY. REV. STAT. ANN. § 158.6455 (Michie/Bobbs-Merrill 1992 & Supp. 1994).

335. *Id.*

336. *Id.*

337. *Id.*

338. *Id.* Distinguished educators are “the state’s most outstanding and highly skilled certified educators” chosen by the State Board of Elementary and Secondary Education. *Id.* § 158.782. These educators work full-time for a designated period at a school in decline or in crisis, by serving as “teaching ambassadors.” *Id.*

339. *Id.* § 158.6455.

340. *See id.* For a number of reasons, including difficulties in determining how many schools may fall into this category and the notion that the system may need some time to stabilize before implementing the most severe sanctions, the portion of KERA relating to schools in crisis will not be implemented until after the 1994–1996 biennium. *See id.*

notified of the right to transfer their children out of the school, and the school must develop an improvement plan.³⁴¹ The school is also eligible for school improvement funds and the services of a distinguished educator.³⁴²

Schools in crisis must develop a school improvement plan and receive a distinguished educator.³⁴³ The distinguished educator must evaluate the staff at the school and within six months make a recommendation to the superintendent regarding the continued employment, transfer, or dismissal of all full and part-time certified staff.³⁴⁴ The principal of the school in crisis must notify all parents of the school's condition and their right to transfer their children to a successful school.³⁴⁵

While the school is the base unit for accountability, equivalent high-stakes accountability requirements apply to staff who are not assigned to a particular school in a local school district.³⁴⁶ Such staff includes instructional staff, locally elected boards of education, and also the superintendents of local school districts.³⁴⁷

4. *Accommodating At-Risk Students in the Accountability System*—(a) *Establishing a roster of students for each school*—For which students will each school be held accountable in a high-stakes accountability system? KERA mandates that all students must be included in the accountability system, including those traditionally considered to be at risk of failure³⁴⁸ and transitory students.³⁴⁹ Schools may have a substantial number of students who attend classes only during a small portion of the school year. Thus, the question of how to establish a roster of students for which the school is to be held accountable is not a trivial matter.

341. *Id.*

342. *Id.*

343. *Id.*

344. *Id.*

345. *Id.*

346. *Id.*

347. *Id.* Unassigned instructional staff will receive rewards and sanctions through a system analogous to the system described for the schools in § 158.6455. *Id.* If a district's performance declines for two consecutive biennial periods, the district will be considered "an education development district" and its board members and superintendent will be removed. *Id.*

348. See *supra* notes 185–91 and accompanying text.

349. See KENTUCKY DEP'T OF EDUC., *supra* note 88, at 106 (holding schools accountable in part for all students enrolled in the Kentucky public school system for at least 100 days).

During the 1992–1993 school year, schools were held accountable for all students enrolled in their schools on the twentieth day of instruction.³⁵⁰ This date was chosen because it would allow schools to know early in the year which students they would be accountable for and allow the Department of Education, its contractor, and the local school districts to track students' whereabouts from this early date until the assessments were administered and reported.³⁵¹ This arrangement proved extremely burdensome on all parties, however, and the difficulties in tracking students from the twentieth day of instruction to the point of assessment made the continuation of that policy unreasonable.³⁵² For example, schools were assessing students who were at the school on the day of assessment but who had attended several schools since the twentieth instructional day.³⁵³

Beginning with the 1993–1994 school year, accountability rosters were established on the day of the assessment.³⁵⁴ Although the new policy is more reasonable logistically, neither policy adequately addresses the major concern of schools. Schools do not want to be held accountable for the score of a student who left on the twenty-first instructional day and became impossible to track. Nor do they want to be accountable for the student who started school on the day prior to the beginning of testing. Although these two conditions represent the extremes, they illustrate the problem of defining an accountability roster that includes all students.

b. Students with disabilities—At-risk populations include students with disabilities. These students must be included within the accountability process in either of two ways. First, they may be included in the assessments by adapting the assessment to the student's disability according to the student's Individual Education Plan (IEP) or 504 Plan.³⁵⁵ If students

350. *Id.* During the 1991–1992 school year, this date was delayed until December 15, 1991 because of the timing of contract approval. *See id.*

351. *See id.*

352. *Id.*

353. *See id.*

354. *See id.* However, schools were only held accountable for a student's writing portfolio if that student was enrolled in the school on the day that the portfolios were due and had been enrolled in the Kentucky public school system for at least 100 days. *Id.*

355. IEP and 504 Plans refer to specific education plans developed for students with recognized disabilities. Kentucky Department of Education Program Advisory No. 93-OCAA-104 (Feb. 1993) (on file with the *University of Michigan Journal of Law Reform*). Under federal law, the plans must include teacher and parent input, review, and agreement. *Id.*

cannot participate in the assessments with all available assistance and adaptive devices, they must participate through the Alternate Portfolio program.³⁵⁶ This second program includes fewer than .5% of Kentucky's school-age population.³⁵⁷

Students taking the assessment with instructional adaptations are measured against the same standards applied to all students.³⁵⁸ Adaptations are made to the administration of the assessment, as opposed to the evaluation or scoring of the assessment.³⁵⁹ Those participating in the Alternate Portfolio assessment make the same contribution to an accountability index as do students through the regular assessment.³⁶⁰ Students in the Alternate Portfolio program are evaluated against a curriculum appropriate for students with disabilities that prevent them from participating in the regular curriculum.³⁶¹ The curriculum to which a student is exposed and the goals of that curriculum determine whether a student should participate in the Alternate Portfolio program or in regular assessment with appropriate adaptations.

CONCLUSION

With KERA, Kentucky has undertaken school reform on a massive scale. In addition to the accountability system detailed here, KERA includes several other significant innovations, such as a major technology initiative and ungraded classrooms for the first four years of school.³⁶² This drastic change is unique within the experiences of the Kentucky public education system and probably nationwide.³⁶³

Inequities in financing brought education reform to the attention of Kentucky's courts. However, much of the justification for asking the courts to consider the financial inequities

356. See *supra* notes 223–25 and accompanying text.

357. KENTUCKY DEP'T OF EDUC., *supra* note 88, at 40.

358. See *id.*

359. *Id.*

360. See *id.* at 40. Those participating in the Alternate Portfolio were not included in the first accountability cycle ending with the 1992–1994 biennium because it was not possible to develop and implement the program until the beginning of the 1992–1993 school year. See *id.*

361. *Id.* (targeting twenty-eight academic expectations).

362. See, e.g., KY. OFFICE OF EDUC. ACCOUNTABILITY, *supra* note 3, at 53–58, 147–59.

363. See STEFFY, *supra* note 126, at 257.

of the Kentucky schools came from the inability of the schools to produce an educated, nationally competitive population. Consequently, both the state supreme court and the General Assembly responded by mandating a school system designed to make Kentucky students as well-educated as any in the world. They recognized that, while equity is linked to the distribution of resources, equity cannot be achieved without addressing the quality of the instructional program. In Kentucky, quality of student achievement has been added both as a new dimension within considerations of educational equity and as a condition for increased funding of the public school system.

In the past, Kentucky and other states have implemented assessment programs that were perceived as having high-stakes accountability. This accountability was typically in the form of public reporting and comparisons of district level data within the state.³⁶⁴ KERA has raised these stakes considerably in that schools and school districts may experience substantial financial rewards or sanctions depending on their students' performance.

High-stakes accountability has imposed new pressures and responsibilities on both the assessment and instructional process. Within this environment, student assessment cannot be designed and implemented without confronting the impact that these assessments will have on the daily instruction and on the educational experiences of students. Instruction and assessment can no longer be designed independently of each other. A statistically well-designed assessment conceived to run efficiently in terms of time and financial costs still may be indefensible and too expensive if it causes the instructional process to narrow its curriculum inappropriately. Such an assessment may be too expensive if it results in an instructional program that does not cause teachers to challenge all students to reach for high levels of performance or to apply and communicate their academic achievements. In the end, any assessment and accountability system that puts at risk our ability to produce the quality students we seek is a system that is too expensive.

364. *Id.* at 91, 257.

