DOCUMENT RESUME

ED 474 274 TM 034 814

AUTHOR	Luvisi, Christopher; Miller, Stephen K.
TITLE	The Effects of Kentucky's Primary Program on Three Measures of Academic Achievement.
PUB DATE	2001-04-00
NOTE	49p.; Paper presented at the Annual Meeting of the American Educational Research Association (Seattle, WA, April 10-14, 2001).
PUB TYPE	Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	Academic Achievement; Age Grade Placement; *Educational Change; Multigraded Classes; *Nongraded Instructional Grouping; Primary Education; *State Legislation
IDENTIFIERS	Kentucky; *Kentucky Education Reform Act 1990; Kentucky Instructional Results Information System

ABSTRACT

In Kentucky, nongraded primary education became a reality under the mandate of the Kentucky Education Reform Act (KERA) in 1990. Research has produced some anecdotal information on the degree to which schools have implemented the ungraded primary program, but there has been no research to link the implementation of the primary program to student achievement or school improvement. Under the regulations growing from the KERA, schools devised different organizational arrangements for the ungraded classrooms, characterized by developmentally appropriate practices and multi-age, multi-ability classrooms. The variance in how students were configured in the multi-age classrooms continued to be a source of conflict and discussion regarding the nongraded primary school program. This study examined the relationship, if any, between the degree of implementation of the primary program and three measures of student achievement: (1) the Kentucky Instructional Results Information System (KIRIS) Growth Index Score for cycle 3; (2) the KIRIS percent improvement score; and (3) achievement test scores for grade 3. Data were available for 463 of the state's 813 schools with primary education programs. Data show that Kentucky schools are improving, but it is difficult to determine the effects of the primary program. The degree and nature of implementation varied widely; political pressures and teacher and school interpretations of the nature of the changes worked to blur the impact of the primary program. Additional research is clearly needed to determine the impacts of the ungraded primary program. (Contains 7 tables and 25 references.) (SLD)

THE EFFECTS OF KENTUCKY'S PRIMARY PROGRAM ON THREE MEASURES OF ACADEMIC ACHIEVEMENT

Christopher Luvisi*
Goshen (KY) Elementary School
and
Stephen K. Miller
University of Louisville

Paper presented at the American Educational Research Association Annual Meeting

April 10-14, 2001, Seattle

*Now deceased; address correspondence to the second author:

Dr. Stephen Miller
Dept. of Leadership, Foundations and Human Resource Education
University of Louisville
Louisville, KY 40292
(502) 852-6475
steve.miller@louisville.edu

THE EFFECTS OF KENTUCKY'S PRIMARY PROGRAM ON THREE MEASURES OF ACADEMIC ACHIEVEMENT

Introduction

With every new school reform initiative there is the ultimate goal to understand whether or not schools and their students are better off because of the initiative or are in fact, worse off than they were before school reform arrived. Understanding the impact the statewide implementation of the nongraded primary program in Kentucky has had on student achievement, if any, is the focus of this study. Nongraded elementary school reforms have been around since the mid to late 1950s in the United States, expanded exponentially in the 1960s and 1970s, dwindled in the 1980s, and began a resurgence in the 1990s (Anderson & Pavan, 1993; Goodlad & Anderson, 1987). In Kentucky, the nongraded concept became a reality due to the mandate of the Kentucky Education Reform Act (KERA) in 1990. All elementary schools began implementing the nongraded concept by the 1992-93 school year. Is there any relationship between the implementation of the primary program and school achievement? The research to date has been wide ranging in examining the primary school program but very limited in revealing the relationship, if any, between the primary program, student achievement, and school improvement. While some research has produced anecdotal evidence on the degree to which schools have implemented the primary program, no effort to date has linked the implementation of the primary program to student achievement or school improvement on a wide scale basis.

Kentucky Education Reform Act (KERA)

In 1985, the superintendents in 66 school districts (out of 177) with the lowest per

pupil property values created a group called "The Council for Better Education" and filed suit against the Commonwealth of Kentucky, charging that their public school districts had been denied equal educational opportunities due to the system of financing public schools. The suit challenged the state's funding formula for placing too much emphasis on local resources, resulting in "inadequacies, inequities, and inequalities throughout the state." The Kentucky Supreme Court ruled the state system of education unconstitutional in June 1989 (ERIC Review, 1995). "All statutes creating, implementing, and financing elementary and secondary education were invalidated and the court ordered the General Assembly to form a new system of public schooling" (Lumsden, 1996, p. 4). That same summer the state formed a task force on education reform that included legislators and members of the governor's office, as well as business and education leaders. The task force was divided into three committees: finance, governance, and curriculum. By March 1990, the Kentucky Education Reform Act (KERA) had been adopted by the Kentucky General Assembly and that July signed into law by then Governor Wallace Wilkinson. To create a system of schools that were accountable and performed at a high level, Kentucky embarked on a set of reforms that were described as a "high stakes" accountability system. KERA was based on the beliefs that: (1) left alone, schools will not systematically change; (2) change must be statutorily mandated; (3) assessment drives instruction; (4) sanctions and rewards are required to force change; and (5) sanctions and rewards must be linked to assessment (Steffy, 1993). This is a significant point in this study due to the changes required by elementary schools in regards to the nongraded primary school program.

Kentucky's Nongraded Primary School Program

As described in *The Wonder Years* (Kentucky Department of Education, 1991) "[Kentucky's Primary School Program] is characterized by developmentally appropriate practices, multi-age/multi-ability classrooms, continuous progress, authentic assessment, qualitative reporting methods, professional teamwork, and positive parent involvement" (p. 9). These factors are referred to as the seven critical attributes.

Schools devised various organizational arrangements, which were permissible under the regulation. Elementary schools had the freedom to design multiple-year (combined grade levels of more than two grades) and/or two-year combined classrooms (commonly referred to as dual-age groups) to address the seven critical attributes of the primary program. Ultimately, it was up to each school to determine the degree to which students would participate in the nongraded primary school. This discretion allowed schools to create a wide variety of organizational arrangements creating the context for which the other six critical attributes would be implemented. The variance in how students are configured in multi-age classrooms continues to be a source of conflict and discussion regarding the nongraded primary school program.

The requirement to educate young children in mixed-age groups rather than the traditional graded organization would become the least popular attribute of the primary school program (Winograd, Petrosko, Compton-Hall, & Cantrell, 1997). Thus, the greatest opposition to the primary school program had more to do with the required mixed-age attribute than other instructional practices required by the legislature.

By June 1992, every elementary school in the state had developed an action plan to begin implementation of the primary school program. Teachers worked in teams for weeks and months creating a school plan to implement the seven critical attributes of the

primary school. All schools were required to do away with traditional graded schools for children in kindergarten through third grade initially and form new mixed-age arrangements placing children from different grade levels together in the same classroom. Mixed-age groupings could include children that ranged in age from 5-year-olds to 9-year-olds. As previously noted, 5-year-old students could be exempt from the multi-age/multi-ability component of the regulation if the school thought it appropriate. Currently, the Kentucky Department of Education is reporting that in twenty-five percent of classrooms, 5-year-old students are mixed with older students for part of their day (Miller, 1999). The multi-age/multi-ability attribute of the primary school program was created to ensure that children would be in a classroom environment that allowed all children to progress at their own rate. Some children would need more time to master curriculum than others would, and the multi-age requirement of the primary school was designed to allow this instead of retaining a child and repeating the same experience.

Schools would be held accountable and also have the flexibility to design curriculum to meet the needs of all students. Elementary age students would enter the primary school in kindergarten and exit four years later when they had successfully completed the requirements of the primary school program and were ready for fourth grade. Those students who did not successfully complete the primary program would be given an additional year. The goal was for students to learn in an environment where they could progress at their own pace based on their readiness. Ultimately, the primary school program was designed to provide young children with the skills and knowledge that would foster future success. Multi-age arrangements were part of the design to foster a developmental approach to teaching and learning.

Organization of Primary School

When the Kentucky Department of Education developed documents to guide schools and districts on how to organize primary schools, there was a wide array of options for schools to choose from to implement the nongraded component of the primary school program. Children could be organized differently as regards to class groupings from school to school or within the same school.

The Kentucky Department of Education provided little direction as to which grouping pattern was more effective. The Department of Education funded fourteen Primary Program Resource Schools to pilot the new Primary Program and many more schools began early implementation and experimentation on their own. Many of the teachers who were a part of the resource schools visited nongraded primary schools in Ohio, British Columbia, and other places that had implemented the model based on Goodlad and Anderson's (1987) work. Schools made their own decision as to the nongraded configuration of the primary program. Consequently, many different combinations of grades were implemented across Kentucky. That variance of decision-making is at the core of this investigation. (What may be viewed as "best" for some schools obviously did not seem "best" to others.) The focus of this study is the level of implementation each school reported based on the Innovation Component Configuration Map (see Luvisi, 2000, Appendix A) and the relationship to school achievement. Table 1 illustrates the different combinations.

Table 1

Multi-age Arrangements for Nongraded Primary School Program

Type of Group	Configuration
Multi-year	children in 1st, 2nd, and 3rd years (formerly K, 1, 2)
	children in 2nd, 3rd, and 4th years (formerly 1, 2, 3)
	children in 1st, 2nd, 3rd, and 4th years (formerly K, 1, 2, 3)
Dual-year	children in 1st and 2nd years (formerly K, 1)
	children in 2nd and 3rd years (formerly 1, 2)
	children in 3rd and 4th years (formerly 2, 3)
Single-year	children in the same year of the program in one class

Note. Other configurations are possible under the regulation for primary school. For example, schools may include 4th and/or 5th grade in the primary school when combined with the 4th year of primary school. If schools decided to combine those students, then the regulations for primary school would extend to those grades (Kentucky Department of Education, 1991, p. 57).

Research Questions

The research questions guiding this study address assumptions regarding instructional practices and student achievement in the elementary school setting. The state of Kentucky made a decision to break new ground in the field of education in the hope that new innovations in governance, finance, and instruction would lead to a better education for all school age children. The nongraded primary program was designed as a foundation piece of legislation that would increase the likelihood of school success for thousands of young children in Kentucky who had historically been left behind

educationally. The convergence of practices in early childhood education advocated by the KERA initiative and high stakes accountability was designed to create a climate where all students would learn at high levels. As schools across Kentucky struggled to find the blend of organizational and instructional practices that would lead to higher student achievement, they consciously decided the degree to which they would implement key attributes of the primary school program to leverage higher test scores on the state assessment system. This study examined the relationship, if any, between the degree of implementation of the primary program and three measures of student achievement: KIRIS Growth Index Score for Cycle 3, the KIRIS Percent Improvement score for schools from Cycle 2 to Cycle 3, and CTBS/5 (grade 3) achievement test scores. KIRIS is the Kentucky Instructional Results Information System, the testing structure developed by the state to drive accountability in a value added long term growth model (cf. Miller, 1992).

All three scores provided a different perspective to understanding the nongraded primary and student outcomes. First, the Growth Index reflects six years cumulative achievement through Kentucky's high stakes accountability system (KIRIS) and reflects a score comparable to all other Kentucky elementary schools. This is significant because no state has mandated all its schools to implement a nongraded primary school plan and make it accountable through nationally normed assessments. Second, the Percent Improvement statistic from Cycle 2 to Cycle 3 was analyzed to determine if there is a relationship between schools that are making progress (improved performance) and those that are not after the nongraded primary program has been implemented for six years. A school showing a decline or low growth in the percentage of students viewed as

successful by the state accountability system versus those schools showing steady or high growth could have policy implications as it relates to the implementation of the nongraded primary school program in Kentucky. Finally, the Comprehensive Test of Basic Skills is a nationally normed reference test and provided comparisons for the Total Battery to other student populations nationally. Of particular interest, based on preliminary work of analyzing outlier data related to growth scores (value-added change), there was no obvious factor associated with schools who are at the extreme ends. Thus, this study may bring better understanding of factors that affect these growth relationships, if any.

Given this context, the central question for this study is, Does the degree of implementation of the primary school program in Kentucky affect measures of student achievement? To answer this, three primary questions are investigated:

Based on two critical attributes of the nongraded primary school program--multi-age/multi-ability classrooms and developmentally appropriate practices:

1. How does the degree of implementation affect the school level KIRIS Growth Index score?
2. How does the degree of implementation affect the school KIRIS Percent Improvement score?
3. How does the degree of implementation affect the school CTBS/5 (grade 3) achievement test scores?

For each of the three primary questions, two separate relationships are examined:

a. To what extent are the two critical attributes related to the measures of student achievement?

b. What is the relationship between the critical attributes and the measures of student achievement, holding constant the demographic factors?

Methods

Schools in Study

This study involves those primary schools in the state of Kentucky which provided complete feedback to the Kentucky Department of Education by using the KIER Primary Program Innovation Component Configuration Map entitled "Primary Program Improvement Plan" by May 30, 1997 (see Luvisi, 2000, Appendix A).

During the 1996-97 school year there were (approximately) 813 elementary schools in Kentucky with primary education programs (Kentucky Department of Education, 1997). Each of those schools received a letter from the Department of Education describing the survey and the process for reporting back to the Department. Of the 813 elementary schools with primary programs identified, 730 reported 1996-97 Primary Program Demographic Survey data to the Division of Primary Education. There were 69 elementary schools with primary programs that did not report data that year. Fourteen elementary schools could not be accounted for in a cross reference of the School Directory list provided by the Division of Assessment Implementation and the elementary schools that were identified with primary programs. Thus of these 813 during the 1996-97 school year, 463 schools (57%) provided data to the Kentucky Department of Education that were usable for this study.

The population of 463 schools primarily represents rural and small town elementary schools across seven of the eight regional service center areas defined by the Kentucky Department of Education. No data were used from Region 3 because they

were unavailable. A combination of parents, administrators, teachers, and members of the Site Based Decision-Making (SBDM) Council completed the survey at each school.

Survey Instrument

The Innovation Configuration Map for Primary Programs was developed by the University of Kentucky, Institute on Education Reform in collaboration with KIER (Bridge, 1994). The configuration map was originally designed as a research tool to assess a primary school or classroom setting to determine the level at which the school or classroom was implementing the constructs of the primary school program. While the document was developed as a research tool, it was also designed for use as a planning instrument and self-assessment of local patterns of implementation of the primary program.

The Kentucky Department of Education used the instrument in the spring of 1997 for schools to self-assess their level of implementation of the primary school program. The Innovation Component Configuration Map For Primary Programs has five different areas for assessment purposes: a) Learning Environment, b) Developmentally Appropriate Practices, c) Assessment, d) Educational Partnerships, and e) Multi-Age/Multi-Ability Grouping Patterns. The configuration map consists of 48 different characteristics of a primary school program across the five areas above (see Luvisi, 2000, Appendix B). Only section 2, Developmentally Appropriate Practices (16 key components) and section 5, Multi-Age/Multi-Ability Grouping Patterns (5 key components) were used in this study. Each component has four different variations: a, b, c, and d, printed from left to right on pages of the map. Those variations farthest to the left (a) best reflect the emerging practice advocated by the KERA initiative according to

its developers.

For example, item B. 2) of Developmentally Appropriate Practices describes the *balance of student and teacher initiation.* At the highest level of implementation a school team would circle (a) which states, "There is extensive evidence of both student and teacher initiated activities"; or (b) the next highest level of implementation, "There is extensive evidence of teacher initiated activities with some opportunities for student initiated activities"; or (c) the third highest level of implementation, "There are few student initiated activities"; or (d) which is the lowest level of implementation, "There are no opportunities for students to initiate activities."

In section 5 of the Configuration Map, Multi-Age/Multi-ability Grouping Patterns, key component 2, *Years with the same teacher*, the item considers the variation of emerging practice as advocated by the KERA initiative: (a) states that children stay with the same teacher for two or more years while variation (d) states that children do not remain with the same teacher for more than one year. Variations (b) and (c) have more permissive language such as "may," which describes practices between (a) and (d), in effect a continuum ranging from most to least advocated. In short, those schools choosing variation (a) are reporting practices advocated by the KERA initiative while those reporting variation (d) for any component are reporting practices not advocated as an emerging practice of the KERA initiative, according to practitioners, researchers, and developers of the Innovation Component Configuration Map for Primary Programs.

Role Group Respondents

An evaluation team from each school completed the "Primary Program Improvement Plan." The team was comprised of one or more of the following groups:

parents, administrators, teachers, and SBDM council members. Of the 428 schools reporting evaluation team data from the population of 463 schools used in this study, 96% of the schools used two or more of the above groups to form their evaluation team to report implementation data to the department of education. Sixty-two percent of schools used parents (1,021 total parent team members), 86% of schools used administrators (370 total administrative team members), 98% of schools used teachers (3,677 total teacher team members), and 78% of schools used SBDM council members to form evaluation teams. Of the 333 total SBDM team members, it is likely that some of these overlap with the other role groups above.

Other Data

The Kentucky Department of Education provided accountability-based achievement data which included the KIRIS Growth Index score from Cycle 3. A KIRIS Percent Improvement score was also provided by KDE which calculated the percentage each school exceeded or fell short of its baseline score from Cycle 2 (1994-95 and 1995-96) to Cycle 3 (1996-97 and 1997-98). Also included were CTBS/5 total battery (grade 3) normal curve equivalent (NCE) school scores from the 1997-98 school year.

The Kentucky Department of Education also provided the free and reduced lunch data for the 1996-97 school year, size of each school's membership, and regional location of the school. The free and reduced data are the actual data reported by the school district for each school included in this study, based on approved federal applications.

Conceptual Model

The unit of analysis in this study was the school. The main purpose of the research was to examine the predictive power of primary school program variables for

each of the three measures of student achievement. The study included the designation of two types of independent variables, those which served as demographic factors or controls and those that are attributes of the primary school program, along with the dependent variables, which were the three measures of school outcomes. The conceptual model describes the relationship between these variables (see Figure 1 below).

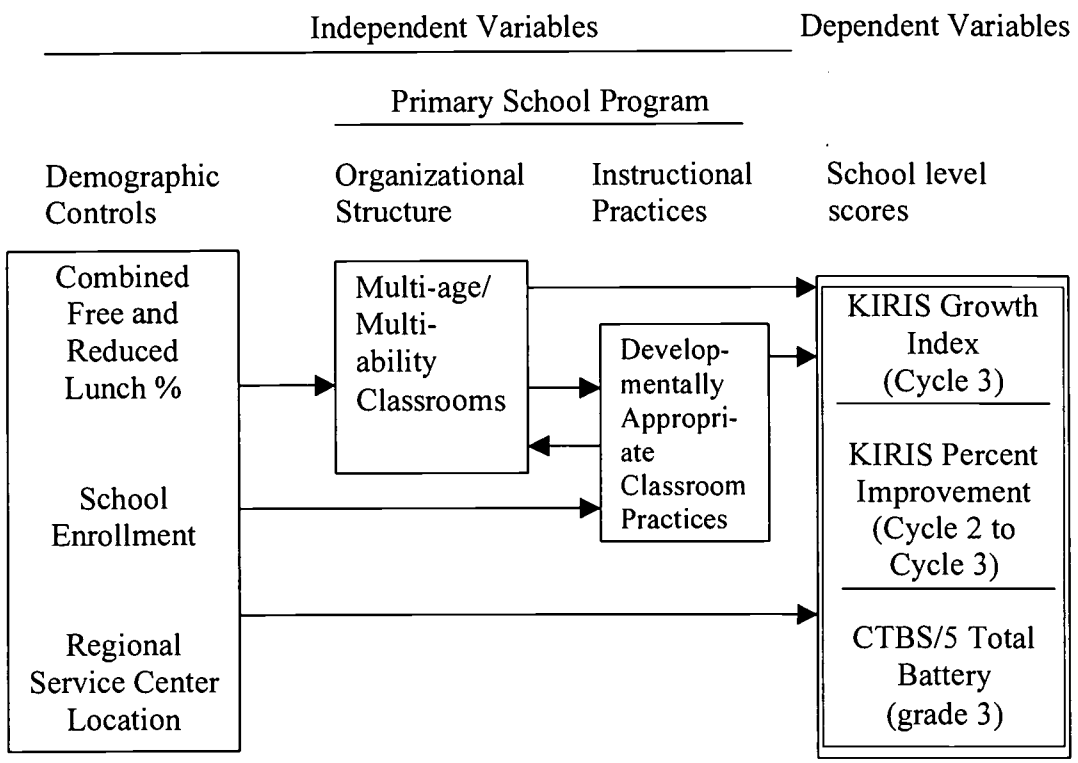

Figure 1. Conceptual model for research design. Demographic controls are used to measure influence on both Primary School Program variables and dependent variables. Primary program variables were examined for influence on each other as well as on dependent variables. For all relationships, the three dependent variables are analyzed separately.

Independent Variables

The demographic variables included those found to be significant to student achievement and geopolitical governance. In this study, the socioeconomic status (SES)

data for each school, combined free and reduced lunch percentage, was obtained at the state level. Race was not used in this study because the schools are primarily small town, rural, with a very low percentage of African-American and other minority groups, which therefore are too underrepresented to make any distinctions. School enrollment and Regional location were likewise obtained from the state.

Primary School Program Variables

The alterable independent variables for this study were two of the seven critical attributes (multi-age/multi-ability classrooms and developmentally appropriate practices) of the primary school program. The innovation configuration map contains five different sections. Two of those sections have a one-to-one correspondence with the critical attributes of the primary school program, which are used in this study. Luvisi (2000, Appendix A) lists the five components for multi-age/multi-ability classrooms and the sixteen key components that define developmentally appropriate practices, according to the developers of the configuration map.

Each of the 21 components required rating an aspect of the primary program on the level of implementation from high (a) to low (d). For the purposes of this study each response was coded as follows: a = 4, b = 3, c = 2, and d = 1. The numerical coding represents the level of implementation a school reported for any one component of the primary program being measured. The raw implementation survey data were entered into SPSS 10.0 and transformed into variables. Once the data were entered, the individual responses for each school were totaled and means were calculated by averaging the five responses for multi-age/multi-ability grouping patterns (MA/MA), resulting in one score for that aspect of the primary school program. The same was done for developmentally

appropriate practices (DAP), the sixteen items describing this component of the primary school resulting in one score for each school. These mean scores for MA/MA and DAP constitute the two Primary Program variables.

The 21survey items for the two Primary Programs attributes were ultimately analyzed by examining simultaneous and hierarchical regressions, respectively, for the two relationships detailed in the research questions. (The calculations were constructed by using school means.) Each school was assigned an official identification code by the Kentucky Department of Education.

Prior to conducting the multiple regressions, psychometric analyses were conducted on the internal consistency of both primary school variables. Since each of the two primary program factors was measured by more than one item, it was necessary to determine the level of internal consistency among the five multi-age/multi-ability items and the sixteen components describing developmentally appropriate school practices. Cronbach's alpha and factor analysis were used to examine internal consistency of the items. The final inclusion of the 5 items and 16 items, respectively, in the two Primary Program factors depended on the results of the psychometric tests (see Luvisi, 2000). The following are complete definitions of each variable.

1. *Multi-age/Multi-ability Grouping Patterns:* There must be evidence of different age levels present in the classroom; students are with the same teacher for multiple years, five-year-old students are with the same teacher for multiple years, five-year-old students are part of a class with older students and participate with students of other age groups for all or part of their instruction, and students with disabilities are fully included in the regular

classroom setting.

2. *Developmentally Appropriate Practices:* There must be evidence of integrated instruction which uses Kentucky's Learning Goals, flexible scheduling, broad-based themes and units, authentic problems, levels of questioning, meaning centered reading, meaning centered writing, problem solving mathematics, discovery science, inquiry-oriented social studies, other subject areas, and varied instructional strategies which utilize balanced instructional delivery, balance of student and teacher initiation, active child involvement, flexible grouping and continuous progress.

Dependent Variables

The three dependent variables are outcome variables that were calculated at the school level. They include the KIRIS Growth Index for 1997-98, KIRIS Percent Improvement, and 1998 Total Battery grade 3 score on the Comprehensive Test of Basic Skills Survey Test Form 5 (CTBS/5).

The definition for the Growth Index comes from the *Kentucky School and District Accountability Results Accountability Cycle 3 (1994-95 to 1997-98) Briefing Packet* (Kentucky Department of Education, 1998a):

> [*KIRIS Growth Index* refers to the] statistic that describes the school or school district's status during the second two years of an accountability cycle. It is compared with the corresponding baseline data. Examples follow: Growth Index 3 is the statistic that results from calculating the Weighted Average of the 1996-97 and 1997-98 Growth Indices. Growth Index 2 is the statistic that results from calculating the Weighted Average of the 1994-95 and 1995-96 Growth Indices.

Growth Index 1 is the statistic that results from taking the Weighted Average of the 1992-93 and 1993-94 Growth Indices. (p. 32)

The definition for Percent Improvement is found in the Kentucky Department of Education (1998b) report, *Technical Considerations In Using The December 3, 1998 Release of Kentucky Accountability Results For Schools and Districts Accountability Cycle 3 Results (1994-1995 to 1997-1998):*

[*KIRIS Percent Improvement* refers to the statistic which] compares the school's Combined Growth Index 1996-1998 and Baseline 1994-1996.

Formula used is:

(GROW9798 – BASELINE) / BASELINE * 100.

Result is rounded to two decimal places and can be negative. (p. 6)

The Comprehensive Test of Basic Skills Survey test Form 5 is a national norm-referenced test. The reported score (CTBS/5 grade 3) is the total battery score interpreted as a Normal Curve Equivalent (NCE) (Kentucky Department of Education, 2001, p, 14).

Data Analysis

The procedures for data analysis were consistent with each research question. The survey data were calculated into descriptive statistics using SPSS 10.0; then multiple regressions were run as the principal method of data to answer the research questions. Multiple regression is based on the prediction of a dependent variable from a set of predictors (Stevens, 1992). The set of predictors for this study was constructed from both demographic variables (non-alterable statistical controls) and the two Primary Program variables (alterable independent variables) (cf. Bloom, 1980). The dependent variables are the three separate school level scores that measure academic achievement and growth.

For the first relationship (a), the two primary program variables were simultaneously entered into a regression equation that has the school score as the dependent variable; this equation is repeated for each of the three dependent variables. For the second relationship (b), entering first the demographic variables into the equation as controls and then the primary program variables was done to examine the net effect of the primary program components on the school scores. Again, this procedure was conducted for each of three different school scores.

Results

The first relationship examines whether the two critical attributes of the Primary Program, Multi-Age/Multi-Ability (MA/MA) and Developmentally Appropriate Practices (DAP), affect student outcomes. Tables 2 through 4 summarize these data; the two primary program variables were simultaneously entered into the three separate regression equations for the three measures of the school score as the dependent variable.

For Table 2 there is no statistically significant relationship for KIRIS by MA/MA and DAP. Multi-Age/Multi-Ability and developmentally appropriate practices are not predictors of KIRIS Growth Index scores. The overall ANOVA has an F ratio of 2.878 with Sig. $F = .057$, just beyond the $p \leq .05$ decision rule for this study. While not significant overall, the t values for MA/MA and DAP indicate that the latter has the stronger relation to KIRIS achievement. This may be relevant for subsequent studies because the overall ANOVA is so close to being significant.

Table 2

Simultaneous Regression for KIRIS Growth Index Scores on Primary Program Components

Model fit	
Multiple *R*	.111
R Square	.012
Adjusted *R* Square	.008
Standard Error	6.7546

Analysis of Variance			
	Sum of Squares	*df*	Mean Square
Regression	262.624	2	131.312
Residual	20987.632	460	45.625
Total	21250.257	462	
$F = 2.878$	Sig. $F = .057$		

Variables in equation					
Variable	*B*	*SE B*	Beta	*t*	Sig. *t*
MA/MA	-.424	.619	-.034	-.685	.494
DAP	2.625	1.095	.117	2.398	.017

Table 3

Simultaneous Regression for KIRIS Percent Improvement Scores on Primary Program Components

Model fit	
Multiple *R*	.087
R Square	.008
Adjusted *R* Square	.003
Standard Error	11.5797

Analysis of Variance			
	Sum of Squares	*df*	Mean Square
Regression	466.214	2	233.107
Residual	61681.147	460	134.089
Total	62147.361	462	
$F = 1.738$	Sig. $F = .177$		

Variables in Equation					
Variable	*B*	*SE B*	Beta	*t*	Sig. *t*
MA/MA	-1.427	1.061	-.066	-1.345	.179
DAP	-1.489	1.877	-.039	-.794	.428

In Table 3, the primary program attributes are examined as predictors of the change score, KIRIS Percent Improvement (PCTIMP). Both multi-age/multi-ability and developmentally appropriate practices demonstrate no statistically significant relationship

to predicting the value-added score from Cycle 2 (1994-96) to Cycle 3 (1996-98) (overall ANOVA, $F = 1.738$, Sig. $F = .177$).

Table 4

Simultaneous Regression for CTBS/5 Scores on Primary Program Components

Model fit	
Multiple *R*	.162
R Square	.026
Adjusted *R* Square	.022
Standard Error	7.5652

Analysis of Variance			
	Sum of Squares	*df*	Mean Square
Regression	711.164	2	355.582
	Sum of Squares	*df*	Mean Square
Residual	26327.093	460	57.233
Total	27038.257	462	
$F = 6.213$	Sig. $F = .002$		

Variables in Equation

Variable	*B*	*SE B*	Beta	*t*	Sig. *t*
MA/MA	-2.356	.693	-.165	-3.398	.001
DAP	2.424	1.226	.096	1.977	.049

Table 4 summarizes the regression analysis when CTBS/5 (grade 3) is the dependent variable and the two primary program components are independent variables. The overall ANOVA is significant ($F = 6.213$, Sig. $F = .002$). Both primary program attributes are statistically significant predictors for CTBS grade 3 scores, although they only account for 2.2% of the variance. The regression coefficient for multi-age/multi-ability is statistically significant at $p = .001$ while the coefficient for developmentally appropriate practices is significant at $p = .049$. The relation is negative for MA/MA with schools which report lower degrees of implementation of multi-age/multi-ability having somewhat higher CTBS scores, although as stated above, little variance can be attributed to either primary program component.

Tables 5 through 7 summarize the data to provide the answer for the hierarchical regression analysis. What is the relationship between the critical attributes and the measures of student achievement, holding constant the demographic factors? For this equation, demographic variables were entered into the equation first and then primary program attributes variables. The school scores were used as the dependent variable, with three separate analyses.

For Table 5, when KIRIS Growth Index is used as the dependent variable in the hierarchical regression, combined free and reduced lunch percentage has a statistically significant negative relationship to KIRIS. There is also statistical significance for one of the regions (V6 contrast). In Model 1, demographic variables are used as independent variables and account for 32.8% of the variance. In Model 2 when the primary program components are added to the equation and the demographic factors function as controls, neither is statistically significant and the amount of variance explained increases only .05

to .333. Free and reduced lunch is the primary predictor even in Model 2 (standardized beta weight of -.561).

In contrast to the regression of KIRIS Growth Index scores in Table 5, when PCTIMP (change) scores are used as the dependent variable in the hierarchical regression (Table 6), there is no significant relationship for any of the demographic factors with the exception of one region (V2 contrast). The percentage improvement from one assessment cycle to the next was not strongly related to the predictors. In Model 1 demographic factors account for only 3.5% of the variance; when primary program components are added to the equation in Model 2, the R^2 added is only .003. Neither primary program variable approaches significance. In summary, the percentage improvement is not well explained in this study, other than to conclude that demographic factors and primary program components have little relationship with growth scores.

Table 5

Hierarchical Regression for KIRIS Growth Index Scores against Primary Program Components, Controlling for Demographic Variables

	Model Fit	
	Model 1	Model 2
Multiple *R*	.583	.590
R Square	.340	.348
Adjusted *R* Square	.328	.333
Standard Error	5.5579	5.5387

Analysis of Variance

	Sum of Squares		*df*		Mean Square	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Regression	7164.111	7321.567	8	10	895.514	732.157
Residual	13900.826	13743.370	450	448	30.891	30.677
Total	21064.937	21064.937	458	458		
F =	28.990	23.867				
Sig. *F* <	.001	.001				

Variables in equation, Model 1

Variable	*B*	*SE B*	Beta	*t*	Sig. *t*
ENROLL	-2.925E-03	.002	-.076	-1.842	.066
FRL	-.183	.015	-.561	-12.018	< .001

(*continued*)

Table 5 (*continued*)

Hierarchical Regression for KIRIS Growth Index Scores against Primary Program Components

Variable	*B*	*SE B*	Beta	*t*	Sig. *T*
V1	-.537	.630	-.044	-.852	.395
V2	.717	.705	.054	1.018	.309
V3	.218	.654	.018	.334	.739
V4	-.619	.653	-.049	-.947	.344
V5	-.164	.641	-.013	-.255	.799
V6	2.306	.699	.178	3.300	.001

Variables in equation, Model 2

Variable	*B*	*SE B*	Beta	*t*	Sig. *t*
ENROLL	-2.856E-03	.002	-.074	-1.803	.072
FRL	-.183	.015	-.561	-11.979	< .001
V1	-.424	.632	-.034	-.672	.502
V2	.572	.706	.043	.810	.418
V3	.335	.654	.028	.512	.609
V4	-.707	.656	-.056	-1.079	.281
V5	5.756E-02	.644	-.005	-.089	.929
V6	2.261	.697	.175	3.243	.001
MA/MA	.604	.527	.048	1.146	.252
DAP	1.350	.916	.060	1.474	.141

Table 6

Hierarchical Regression for KIRIS Percent Improvement Scores against Primary Program Components, Controlling for Demographic Variables

	Model fit	
	Model 1	Model 2
Multiple *R*	.229	.242
R Square	.052	.059
Adjusted *R* Square	.035	.038
Standard Error	11.4137	11.4000

Analysis of Variance

	Sum of Squares		*df*		Mean Square	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Regression	3230.176	3631.103	8	10	403.772	363.110
Residual	58623.047	58222.120	450	448	130.273	129.960
Total	61853.223	61853.223	458	458		
F =	3.099	2.794				
Sig. *F* =	.002	.002				

Variables in equation, Model 1

Variable	*B*	*SE B*	Beta	*t*	Sig. *t*
ENROLL	1.604E-03	.003	.024	.492	.623
FRL	-4.261E-03	.031	-.008	-.136	.892

(*continued*)

Table 6 (*continued*)

Hierarchical Regression for KIRIS Percent Improvement Scores against Primary Program Components, Controlling for Demographic Variables

Variable	*B*	*SE B*	Beta	*t*	Sig. *t*
V1	.525	1.394	.025	.406	.685
V2	3.278	1.447	.145	2.265	.024
V3	2.809	1.342	.135	2.092	.037
V4	-.604	1.341	-.028	-.450	.653
V5	-.283	1.316	-.013	-.215	.830
V6	-.684	1.435	-.031	-.477	.634

Variables in equation, Model 2

Variable	*B*	*SE B*	Beta	*t*	Sig. *t*
ENROLL	1.507E-03	.003	.023	.462	.644
FRL	-4.350E-03	.031	-.008	-.138	.890
V1	.360	1.300	.017	.277	.782
V2	3.519	1.453	.156	2.421	.016
V3	2.615	1.346	.126	1.943	.053
V4	-.480	1.349	-.022	-.356	.722
V5	-.432	1.326	-.021	-.326	.745
V6	-.622	1.435	-.028	-.433	.665
MA/MA	-.839	1.085	-.039	-.773	.440
DAP	-2.336	1.885	-.061	-1.239	.216

Table 7

Hierarchical Regression for CTBS/5 Scores against Primary Program Components, Controlling for Demographic Variables

	Model fit	
	Model 1	Model 2
Multiple *R*	.596	.601
R Square	.355	.361
Adjusted *R* Square	.344	.346
Standard Error	6.1588	6.1457

Analysis of Variance

	Sum of Squares		*df*		Mean Square	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Regression	9400.085	9548.231	8	10	1175.011	954.823
Residual	17068.981	16920.836	450	448	37.931	37.770
Total	26469.066	26469.066	458	458		
F =	30.978	25.280				
Sig. *F* <	.001	.001				

Variables in equation, Model 1

Variable	*B*	*SE B*	Beta	*t*	Sig. *t*
ENROLL	3.404E-03	.002	.079	1.934	.054
FRL	-.231	.017	-.631	-13.688	< .001

(*continued*)

Table 7 (*continued*)

Hierarchical Regression for CTBS/5 Scores against Primary Program Components Controlling for Demographic Variables

Variable	*B*	*SE B*	Beta	*t*	Sig. *T*
V1	-.901	.698	-.065	-1.290	.198
V2	.873	.781	.059	1.118	.264
V3	-2.919	.724	-.215	-4.030	< .001
V4	-1.379	.724	-.098	-1.905	.057
V5	.753	.710	.055	1.061	.289
V6	3.120	.774	.215	4.030	< .001

Variables in equation, Model 2

Variable	*B*	*SE B*	Beta	*t*	Sig. *T*
ENROLL	3.272E-03	.002	.076	1.861	.063
FRL	-.227	.017	-.620	-13.368	< .001
V1	-1.050	.701	-.076	-1.498	.135
V2	.851	.783	.058	1.087	.278
V3	-2.908	.726	-.214	-4.007	< .001
V4	-1.210	.727	-.086	-1.664	.097
V5	-.568	.715	.041	.794	.428
V6	3.197	.774	.220	4.133	< .001
MA/MA	-1.151	.585	-.081	-1.969	.050
DAP	.861	1.016	.034	.847	.398

In Table 7, CTBS/5 scores are used as the dependent variables in the hierarchical regression equation with demographic factors and primary program components used as independent variables. In Model 1, demographic factors account for 34.4% of the variance in CTBS scores. Combined free and reduced lunch percentage has a statistically significant negative relationship for these grade 3 standardized achievement scores. As the percentage of poverty rises as measured by FRL, CTBS scores go lower. Two of the regions were also significant (V3 and V6) although in different directions. In Model 2, there is essentially no difference in the regression when the two primary program components are added to the equation. Although the multi-age/multi-ability primary component is statistically significant ($p = .05$), the amount of variance increases only .02% and the statistically significant controls remain the same. The predictive relationship for MA/MA is negative; schools with higher CTBS scores tend to report more traditional graded configurations. However, the minimal change in the Adjusted *R* Square from Model 1 to Model 2, as noted above, suggests that only the demographic factors have much effect on standardized achievement scores, and that primarily from FRL (standardized beta weight of -.620).

Discussion

For the first relationship (simultaneous regression), the three separate regressions produced three statistically significant results, although the Adjusted R^2 was virtually negligible in each instance. For the KIRIS Growth Index, DAP was significant, but the effect size was only .008 essentially no explanation. Neither of the primary program attributes approached significance for the KIRIS change score. For the CTBS grade 3 scores, both components were significant but the Adjusted R^2 was still very small, now

2.2%.

Of note is the fact that the significant relationship for multi-age/multi-ability with CTBS achievement is negative. One possible explanation is that schools wanting to improve on the state assessments decided to abandon the multi-age requirement and to revert to traditional grouping methods as a way to "teach to the test" more effectively. The Appalachian Educational Laboratory (AEL) (1998) documented the growing lack of support for multi-age/multi-ability grouping. The negative relationship between MA/MA and standardized norm-referenced test scores may also imply that, to do well on the third grade CTBS, educators believe that traditional methods of grouping students for direct instruction are needed so they can perform better on the test. This may well be influencing decisions at the local level on the degree to which schools implement multi-age/multi-ability grouping patterns. This is a significant finding and may contradict earlier research by Veenman (1995) and Guitierez and Slavin (1992) that nongraded approaches to elementary education achievement levels are as high or higher than traditional graded approaches. These findings indicate that schools which multi-age at a higher level compared to other schools may have lower standardized achievement scores. It must also be noted here that demographic variables were not part of this equation and therefore were not controlled.

For the CTBS achievement, developmentally appropriate practices were also significant, but in a positive direction. No research to date had described the relationship between DAP and standardized tests of school achievement. Prior research has focused on organizational arrangements rather than what is happening in the classroom at an instructional level. Given that demographic variables were not controlled in this

regression, the finding is still significant: a correlation exists between developmentally appropriate practices and higher achievement levels on the CTBS/5 comprehensive test of basic skills.

In the second relationship, the same question was asked as in the first except that demographic variables were used as control variables in the hierarchical regression analysis. Neither primary program attribute had any significant variance accounted for in the regression analysis. The finding from the first relationship did not hold up once the demographic factors were controlled.

Implications for Policy

The decision to mandate the nongraded primary program in every public elementary school across Kentucky, replacing grades K-3, was one of the most controversial and often debated pieces of KERA (Kannapel, 1995). The impact of the decision to mandate schools to accept new grouping patterns and instructional practices without the support of teachers, parents, and administrators may have created the greatest educational debate ever noted in Kentucky. As the Kentucky Department of Education scrambled to implement a strategy for the primary program, the forces to subvert and fight the mandate began to gather momentum. It would take two legislative sessions for opponents of the multi-age attribute essentially to undo the requirement, which some schools openly did not implement in the first place. The Kentucky Education Association (KEA) and local Parent Teacher Association (PTA) worked hard to overturn the mandate based on membership concerns and lack of support by parent-teacher groups.

Given the unwillingness by KDE to enforce the implementation of the primary program and the open-ended interpretations of the guidelines themselves, the state may

investigate the need to make the entire primary program optional. Those schools which have embraced the primary program as a grouping and instructional vehicle will continue to do so while other schools which are "primary" in name only can identify themselves with more traditional grouping and instructional practices that reflect separate beliefs without worry. Under these circumstances researchers can better understand the predictive power of the primary program on student achievement since the variance in philosophy should produce concomitantly different grouping practices. These differences in turn would be more likely to be reflected in variance in student outcomes, which subsequent study could document.

From 1993-1998 KIRIS Growth Index scores rose 13.1 points overall (35.7-48.8). Reading scores rose from 32.4 to 58.4 points, math increased 22.1 points (22.3-44.4), science scores improved from 18.2 to 37.2, social studies rose 10.3 points overall (27.6-37.9), and writing improved 7.4 points from a baseline of 31.0 in 1993 (Kentucky Department of Education, 2000). Furthermore, National Assessment of Educational Progress (NAEP) scores in reading increased 3 points from 1992 to 1998 (3 points above the national mean) and math scores improved 5 points over the same time period, although still 2 points below the national mean (KDE). The CTBS/5 was not given statewide until 1997, but elementary schools have already shown an overall improvement of 2 percentiles (50-52) on the Total Battery (KDE). Clearly, schools are improving and the reforms KERA mandated are likely to have had some impact on improving the achievement level of Kentucky students. These improvements include younger students so the Primary Program has obviously been part of this growth.

The implementation of developmentally appropriate practices was statistically

significant and positively related to CTBS/5 scores when demographic factors were not controlled. Pearson correlations also showed positive relationships between DAP and both KIRIS Growth Index and CTBS scores, although not statistically significant for KIRIS. While this is "good news" for KERA advocates, the relatively weak relationship between DAP and school scores is sobering. One possibility for the weak relationship is that many training activities to support teachers focused on "activities" instead of developmentally appropriate instructional techniques (best practices) from an early childhood point of view. While many teachers were enthusiastic about the primary school, they were also under tremendous pressure to "look" like a primary class before they may have been fully ready to implement the attributes of a nongraded concept.

Many primary teachers were trained on how to develop units of instruction on bears, lollipops, pumpkins, and other irrelevant instructional themes instead of learning the underlying strategies of student engagement that are the focus of sound early childhood programs. Teachers came away from training sessions with activities to conduct with their students but many did not understand or support why they were to change teaching practices. Therefore, implementation was often shallow and short-lived and support was tenuous due to the political climate surrounding multi-age grouping strategies. Site Based Decision-Making Councils were to provide the vision and agenda for the primary program through professional development of teachers. Councils, dominated by teachers, would ultimately determine the degree to which the primary program would be implemented. The process of bringing about organizational changes on the school level, as well as instructional and curricular changes, would prove to be a daunting and too often uncompleted task (AEL, 1998).

In another example of restructuring and changing the culture of schools, Keedy (1995) studied three high schools which were recognized as leaders in school restructuring. He found that two of the three schools were no different than what would typically be found in a suburban area even though the schools had a reputation for being student centered. Teachers by and large had not changed teaching practices even though student-centered teaching and learning were a major focus of their restructuring. Changing large organizations or practices on a school and certainly statewide level may be a more daunting task than the original KERA reformers had imagined.

In the primary reform, teachers debated without end what degree of implementation was the most effective or whether to do it at all (Kannapel, 1995). There was little debate, however, concerning the use of developmentally appropriate practices. Teachers generally accepted the notion that children learn at different paces, with different styles, and needs. Many teachers were happy to leave the confines of basal readers and workbooks to embrace integrated studies, inquiry learning, and whole language strategies, but the depth to which teachers understood how to implement developmentally appropriate practices remains unknown. It is clear, however, that teachers believed that they were implementing DAP at a high level. This may have contributed to the high self-reported implementation scores for developmentally appropriate practices. While it is generally accepted that self-report data from surveys such as these tend to be over-reported, surveys which are mailed, rather than telephone or person to person interviews, are generally more accurate (Dillman, 1978). Qualitative measures such as observation over time, in combination with surveys, may lead to a more accurate reading on the relationship between the level of implementation of the primary

program and student achievement. These data need to be captured to understand better how teachers perceive the importance of the primary program as it relates to school achievement.

Leadership at KDE frequently changed and was overly sensitive to public criticism. This was due in part because of the lack of vision and articulation about the most effective way to implement the primary program. Consultants at KDE would provide schools with contradictory information or incorrect information leading to the confusion at the district and school level. (This was observed by the first author who served as a consultant for the Primary School Program.) The primary program was rushed into implementation early in the reform and forced on teachers. More important, as the primary program began to lose its mandate, the accountability system was in high gear.

KIRIS may have played a more important role in diminishing the impact of the primary program from a state policy level. With schools being required to improve or be punished by the state, there was an urgent need to focus on raising test scores as quickly as possible. Strategies such as the four-column method and other short cuts to improving scores on KIRIS focused a greater amount of attention on improving test scores rather than providing a developmentally appropriate setting for young children. The four-column method involves taking the open-response question and breaking it down into smaller pieces so the student understands exactly what the question is asking prior to writing the answer. This is essentially test taking "saavy" and can be a useful equity based strategy (overcoming differences in test taking ability, of which at-risk students are most likely deficient). But these testing strategies are a short term fix. The state policy

to hold schools accountable for student achievement focused the attention on test-driven accountability improvements in the immediate cycle rather than on allowing young children to learn at their own pace. Any connection between appropriate developmental instruction and longer-term achievement growth was thus truncated before being adequately tested.

Further Research

As with all studies, this research suggests a number of follow-up investigations. Some of these are noted here.

The instrument used to assess the level of implementation of the primary program needs to reflect better the practices and realities of implementing such an enormous mandate. Content and pedagogical issues need to be updated to reflect the changes occurring at the school and state levels. The University of Kentucky and KIER collaboratively developed the configuration map. The instrument, An Innovation Component Configuration Map for Primary Programs, had some design flaws, including language which may encourage respondent schools to "over report" what is actually happening at the school level. The instrument, although designed to be a self-assessment tool, had never been used on a broad scale prior to its use by KDE to gather large amounts of data on the implementation of the primary program across the state. Much can be said critically about the instrument, but the large number of schools in this study (463) allowed the authors to assess the data psychometrically. That is an analysis worth doing for the entire instrument (all five sub-components), but for this study, only the two attributes for MA/MA and DAP were examined.

For these two attributes, DAP had satisfactory psychometric properties.

However, the factor of concern was that the 5 items in the MA/MA component as defined by the University of Kentucky (UK) and KIER and as used by the state, produced a low (.62) alpha. Because the focus of the study was the MA/MA attribute as a whole, the attribute was utilized as a single construct. But analysis revealed that one item, on inclusion of special needs students, was not consistent with the other four. A separate study would be needed to examine the remaining 4 components and the special needs item separately.

For that matter, it is not at all clear from the current study how the Primary Program affected special needs students. As indicated above, the school team that responded to the ICC Map interpolated an overall school level. What that "school level" estimate means for special needs students who experience a wide variety of inclusion and resource room models cannot be stated from this study. Further investigation in this area is clearly warranted.

The DAP component as developed by KIER and UK may not be sensitive to the practices that lead to higher achievement scores (i.e., the items are too rough or too broad to capture the fine-grained differences that really count). Researchers, primary teachers, and primary program consultants were on the team that developed the configuration map. Advocacy for particular practices included in the instrument may represent a point of view of the developers rather than sound research-based methodology. Therefore, the components and their variations represent a mix of practices advocated by KERA and the consensus of the developers and may not be directly linked to practices leading to higher achievement gains by students. Further research may include an examination of the beliefs and attitudes of researchers and practitioners to create an instrument which

reflects consensus among the groups so implementation at the school level can be measured more effectively.

In addition, the process by which a school self-reports survey response needs to be consistent for reliability purposes. While 96% of the schools used multiple role groups to complete the self-assessment, those role groups varied from school to school in their composition. Also, the instrument used needs to be redesigned to reflect better the practices across the entire school. The instrument was designed for dual purposes (self-assessment and outside observer) but may not effectively communicate what happened schoolwide. As much variance as there is between schools, the variance may be even greater within schools. The instrument required the respondents to give a holistic picture of the entire school instead of looking at one specific classroom. If the instrument were specifically designed to capture the essence of the mean practices at the school level based on separate classroom practices, it may provide a more accurate snapshot of what is occurring. Future research with the Primary map should include inter-rater reliability studies. These will provide information about how much the ratings on the map are affected by the observer using the instrument. This may lead researchers to understand better how and why the primary program is being implemented at the school level. And on this issue, it would also seem worthwhile for an instrument that could be utilized at the classroom level, as ultimately it is not the mean school level, but the actual classroom level of implementation that affects the students. Clearly work is needed on this classroom instrument so that its data could be used at the level of the individual classroom or aggregated to the level of the school.

There is a great need to know if traditional graded grouping practices rather than

the multi-grade practices advocated by the KERA initiative may lead to better student outcomes and what are some of the consequential benefits of either. Early research (Goodlad & Anderson, 1987; Gutierrez & Slavin, 1992; Pavan, 1992; Veenman, 1995) showed that multi-age classrooms did as well or better than traditional graded arrangements based on academic measures. The research also showed that students in multi-age classrooms had a significant advantage in social and emotional wellness when compared to students in traditional settings. Why is that? This study examined only cognitive outcomes, albeit three different types. But more research needs to be done in Kentucky to determine whether these affective factors are also affected by the use and implementation benefits of multi-age arrangements, as well as by developmentally appropriate practices.

A follow-up to this study should include a look at outliers to determine specific practices at both ends of the achievement spectrum. One suggestion would be to select a set of schools matching low achievers and high achievers. Only schools that fell within a certain range of free and reduced lunch percentage would be included. This would allow the study to focus on the variance caused by factors other than poverty, in effect a set of schools with similar demographic factors but different achievement outcomes. What are the characteristics of these schools with respect to the multi-age component and developmentally appropriate practices? Are these characteristics predictors of school success?

For example, are the different items in the MA/MA component and in the DAP component equally efficacious with respect to producing higher achievement? The current study, by collapsing the 5 and 16 items of the MA/MA and DAP components into

a single construct made the assumption that these items were equally related to achievement. But that may not be the case at all. A different study would be needed to examine the separate contributions of the various items to achievement measures.

Further, by comparing schools at opposite ends of the achievement spectrum, it may be possible to develop a better understanding of the critical factors influencing student achievement. Some schools with high levels of poverty have been successful when measured on the basis of school scores alone, but they are few in number compared to the large number of schools not able to find success at any level due to the level of poverty in their community. Thirty-five high poverty elementary schools (> 50% free/reduced lunch percentage) were identified by KDE as gaining from 10-22.4 points on the CTBS/5 Total Battery from 1997-99 (Kentucky Department of Education, 2000). Total Battery scores ranged from 22-49.9 in 1997 and from 44.1-62.6 in 1999 for the same schools.

These high poverty schools identified as big gainers represent approximately 5% of all elementary schools. Why did these schools improve their scores while others with similar levels of poverty did not? Examination of these schools in terms of Primary Program critical attributes would parallel the approach that the school effectiveness field followed back in the 1970s and 1980s, but would be focused only on the structural/grouping patterns and instructional practices (MA/MA and DAP), or more broadly, might examine all of the critical components of the Primary Program.

The outlier study suggested above typically utilizes high and low scoring schools on stable or absolute measures of achievement such as KIRIS (now CATS) or CTBS scores. But the percentage change score from one accountability cycle to the next also

could play a more important role in measuring the improvement of a school as cited above with the thirty-five high poverty/big gain schools. The focus of the reform act is to provide the resources and accountability measures necessary for all schools to reach a high level of performance over time. The value-added score reflects the continued growth or improvement of a school based on these assessments and represents the KERA mandate for making schools better regardless of the level of poverty. The problem with year-to-year change scores is that they are relatively unstable and unreliable. Thus, outlier studies on high positive and high negative change schools vis-à-vis their implementation of the primary program attributes may prove more useful than the traditional outlier studies. That such study of changing schools is needed is reflected in the current study where significant relationships existed for both of the stable scores (KIRIS Growth Index and CTCBS/5) but essentially no correlates were found for the KIRIS Percent Improvement score.

If so many schools across the state have implemented the primary program at such a high level as reported by the schools themselves, research must be conducted on why there is not a strong relationship between primary program implementation and school achievement. At least six concerns must be considered since the level of implementation was high across the state, as reported by the schools. First, the low variance in responses may "wash-out" any effect on school achievement scores because there are so few differences, leaving other factors to influence the success on school scores. Second, did schools over-report the level of implementation to "look better" for KDE personnel to avoid any unwanted attention from consultants sent to "help"? Third, is there an assessment team composition factor, i.e., do the makeup and number of

teachers, parents, etc., who assess and complete the self-report Primary Program assessment have any relationship to the scores reported? Fourth, has the primary program made a difference with the achievement level of students regardless of demographic variables? Fifth, there is a need to study the effectiveness of the Primary Program training for teachers and administrators. Part of the homogeneous-like responses to implementing developmentally appropriate practices at such a high level could be attributed to teacher perceptions that strategies are being implemented, when in fact they are not. Do teachers and administrators fully understand the depth and breadth of the various components of developmentally appropriate practices, sufficient to report reliably that certain strategies had been implemented? Finally, a separate study looking specifically at the relationships among regions and school scores would be needed to answer more completely the question concerning regional differences and their possible impact on school scores. Researchers who wish to determine whether or not the primary program, as implemented in Kentucky, has any generalizable value within and outside the state must further investigate these concerns.

This study only examined two of the seven critical attributes of the primary program. While it can be argued that these two components, multi-age/multi-ability grouping patterns and developmentally appropriate practices, may be the most important attributes from a structural and behavioral viewpoint, it can also be argued that these two components may best reflect the attitudes of educators implementing the primary program.

But do these two attributes directly capture values and attitudes of teachers? What do educators *believe* about the children they teach every day and what do they believe

about their own efficacy to change the future of these children? Believing children of different age and grade configurations can learn more effectively together in combination with the use of innovative educational practices rather than traditional grouping patterns and practices may be crucial for educators. Furthermore, the seven critical attributes of the primary program *imply but do not directly state* that all children can learn at high levels. Under the traditional, genetic-based, social Darwinist conception of human ability that has dominated American education in the twentieth century, individual differences among students, not how effective are teachers and schools, has been the primary explanation for low achievement among at-risk pupils. Research on these fundamental attitudes and how they relate to teachers' perceptions about and implementation of the primary program is clearly relevant here.

The state of Kentucky has invested 10 years of work in developing the primary school program. To date, there has been no strong relationship established between implementation of the program and school achievement when demographic variables are controlled. Kentucky elementary classrooms have radically changed in the past 10 years in grouping patterns and practices, albeit inconsistently. It is critical that research continue to be done to gain a better understanding on the relationship between the primary program and student achievement. Researchers, administrators, teachers, and parents need to know more about that relationship, specifically factors such as implementation of the primary program, demographic factors (particularly combined free and reduced lunch percentages) and school scores. This understanding is vital to ensure that all children have the same opportunity to learn at high levels.

References

Anderson, R. H., & Pavan, B. N. (1993). *Nongradedness: Helping it to happen.* Lancaster, PA: Technomic.

Appalachian Educational Laboratory. (1998). Evolution of the primary program in six Kentucky schools. *Notes From the Field: Educational Reform in Rural Kentucky, 6*(1), 1-12.

Bloom, B. S. (1980). The new direction in educational research: Alterable variables. *Phi Delta Kappan, 61*, 382-385.

Bridge, C. (1994). *The implementation of Kentucky's Primary Program* (Report of research conducted by Univesity of Kentucky Institute on Education Reform). Frankfort, KY: Kentucky Institute on Education Research.

Dillman, D. (1978). *Mail and telephone surveys: The total design method.* New York: John Wiley & Sons.

ERIC Review. (1995). Kentucky's systemic reform. *The Education Digest, 60*(6), 20-23.

Goodlad, J. I., & Anderson, R. H. (1987). *The non-graded elementary school* (Rev. ed.). New York: Teachers College, Columbia University.

Gutierrez, R., & Slavin, R. E. (1992). Achievement effects of the nongraded elementary school: A best evidence synthesis. *Review of Educational Research, 62,* 333-376.

Kannapel, P. J. (1995). *Opposition to outcome-based education in Kentucky* [On-line]. Available: (ERIC Document Reproduction Service No. ED 390 148)

Keedy, J. (1995). Teacher practical knowledge in restructured high schools. *The*

Journal of Educational Research, 89(2), 76-89.

Kentucky Department of Education. (1991). *Kentucky's primary school: The wonder years*. Kentucky Department of Education, Frankfort.

Kentucky Department of Education. (1997). *Primary program demographic survey data summary*. Frankfort: Author.

Kentucky Department of Education. (1998a). *Kentucky school and district accountability results accountability cycle 3 (1994-95 to1997-98) briefing packet*. Frankfort: Author.

Kentucky Department of Education. (1998b). *Technical considerations in using the December 3, 1998 release of Kentucky accountability results for schools and districts accountability cycle 3 results (1994-1995 to 1997-1998).* Frankfort: Author.

Kentucky Department of Education. (2000). *Results matter.* Frankfort: Author.

Kentucky Department of Education. (2001). *Commonwealth Accountability Testing System, Spring 2001: Interpretive guide* (Kentucky evaluator's ed.). Frankfort, KY: Author.

Lumsden, R. W. (1996). *Kentucky's high stakes accountability system: The explanatory power of school effectiveness variables on secondary schools in decline.* Unpublished doctoral dissertation, University of Louisville, Louisville, KY.

Luvisi, C. L. (2000). *Implementing Kentucky's Primary School Program: Its relation to measures of academic achievement*. Unpublished doctoral dissertation, University of Louisville.

Miller, M. (1999, August). Results of the 1999 primary program demographic survey.

Elementary Forum, 1, 1-2.

Miller, S. K. (1992). Changing conceptions of "good" schools: Implications for reforming urban education. *Education and Urban Society, 25,* 71-84.

Pavan, B. N. (1992). The benefits of nongraded schools. *Educational Leadership, 50*(2), 22-23.

Steffy, B. E. (1993). Top-down-bottom-up: Systemic change in Kentucky. *Educational Leadership, 51*(1), 42-44.

Stevens, J. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Veenman, S. (1995). Cognitive and noncognitive effects of multigrade and multi-age classes: A best-evidence synthesis. *Review of Educational Research, 65*, 319-381.

Winograd, P., Petrosko, J., Compton-Hall, M., & Cantrell, S. (1997). *The effect of KERA on Kentucky's elementary schools: Year one of a proposed five-year study* (Report of research conducted by the University of Kentucky/University of Louisville Joint Center for the Study of Educational Policy). Frankfort: Kentucky Institute for Education Research.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC Educational Resources Information Center

REPRODUCTION RELEASE

(Specific Document)

TM034814

I. DOCUMENT IDENTIFICATION:

Title: The Effects of Kentucky's Primary Program on Three Measures of Academic Achievement

Author(s): Christopher Luvisi and Stephen K. Miller

Corporate Source: Dept. of Leadership, Foundations, and Human Resource Education
University of Louisville
Louisville, KY 40292

Publication Date: April 1, 2001

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

☒

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Stephen K Miller

Printed Name/Position/Title: Stephen K. Miller, Assoc. Prof.

Organization/Address: as above, for Corporate Source

Telephone: 502 852-0637

FAX: 502 852-4563

E-Mail Address: Steve.Miller@louisville.edu

Date: March 17, 2003

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

University of Maryland
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Laboratory
College Park, MD 20742
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2nd Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com