DOCUMENT RESUME

ED 478 672 RC 024 139

AUTHOR Meehan, Merrill L.; Cowley, Kimberly S.; Schumacher, Debbie; Hauser, Brenda; Croom, Nona D. M.

TITLE Classroom Environment, Instructional Resources, and Teaching Differences in High-Performing Kentucky Schools with Achievement Gaps.

INSTITUTION AEL, Inc., Charleston, WV.

SPONS AGENCY Institute of Education Sciences (ED), Washington, DC.; Kentucky State Dept. of Education, Frankfort.

PUB DATE 2003-07-00

NOTE 33p.; Paper presented at the Annual CREATE National Evaluation Institute (12th, Louisville, KY, July 24-26, 2003).

CONTRACT ED-01-CO-0016

PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)

EDRS PRICE EDRS Price MF01/PC02 Plus Postage.

DESCRIPTORS *Academic Achievement; *Classroom Environment; *Educational Practices; *Educational Quality; Elementary Secondary Education; Professional Development; Racial Differences; Relevance (Education); Social Class; Socioeconomic Status; *Teacher Effectiveness; Teacher Expectations of Students; Teaching Styles; Time Management

IDENTIFIERS *Achievement Gap; *Kentucky

ABSTRACT

This study examined differences at the classroom level between Kentucky schools with minimum versus large gaps in academic achievement between particular groups of students. Data were gathered via observations of 213 classrooms at 18 elementary, middle, and high schools. Although all the schools were identified as high-performing in terms of overall academic index scores, nine had minimum achievement gaps and nine had large gaps between socioeconomic or racial groups. Classroom time was used more efficiently and effectively in minimum-gap schools. Teachers in large-gap schools spent more time on administrative routines in their classrooms. Teachers in minimum-gap schools provided more minutes of instruction, more time for student-led activities, and a more appropriate pace. Textbooks and multi-racial materials were used more in minimum-gap schools. The climate in minimum-gap school classrooms was more cheerful, inviting, open, and risk-free, and had less distracting external noises and interruptions. Teachers in minimum-gap schools communicated high expectations to their students, conducted formal or informal assessments of their students, and also provided immediate and corrective feedback to students. The quality of instruction was higher in minimum-gap school classrooms, both overall and for many subscale items. Rural and small-town schools comprised seven of the nine minimum-gap schools and two of the nine large-gap schools. (Contains 29 references) (TD)

Classroom Environment, Instructional Resources, and Teaching Differences in High-Performing Kentucky Schools with Achievement Gaps

Merrill L. Meehan, AEL
Kimberly S. Cowley, AEL
Debbie Schumacher, Kentucky Department of education
Brenda Hauser, Kentucky Department of Education
Nona D.M. Croom, AEL

AEL
P.O. Box 1348
Charleston, WV 25325-1348

Paper Presented at the 12th Annual CREATE National Evaluation Institute
Louisville, KY

July 24-26, 2003

AEL is a catalyst for schools and communities to build lifelong learning systems that harness resources, research, and practical wisdom. AEL serves as the Regional Educational Laboratory (REL) for Kentucky, Tennessee, Virginia, and West Virginia. For these same four states, it operates the Eisenhower Regional Consortium for Mathematics and Science Education. In addition, it serves as the Region IV Comprehensive Center and operates the ERIC Clearinghouse on Rural Education and Small Schools. AEL houses the Institute for the Advancement of Emerging Technologies in Education (IAETE) and the Institute for the Advancement of Research in Education (IARE). The REL contract includes a Technology Specialty for the nation's system of 10 Regional Educational Laboratories.

Information about AEL projects, programs, and services is available by writing, calling, or visiting AEL's Web site.

Post Office Box 1348
Charleston, West Virginia 25325-1348
304-347-0400
800-624-9120
304-347-0487 fax
aelinfo@ael.org
www.ael.org

This publication is based on work sponsored wholly or in part by the Kentucky Department of Education. Its contents do not necessarily reflect the views of the Kentucky Department of Education. This publication also is based on work sponsored in part by the Institute for Educational Sciences, U.S. Department of Education, under contract number ED-01-CO-0016. Its contents do not necessarily reflect the views of IES, the Department, or any agency of the U.S. Government.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

"The term achievement gap is used to denote differences in the academic achievement of particular groups of students" (Reynolds, 2002, p. 4). The achievement gap has been an issue in education for many years. Lucas (as cited by D'Amico, 2001) notes that as early as 1785, Thomas Jefferson saw it as an important issue when he wrote his notes on Virginia. Lucas also points out that W.E.B. Du Bois made its elimination a cornerstone of his agenda (as cited by D'Amico, 2001). According to Viadero (2000), the gap has been documented since the 1960s, at least.

An achievement gap is seen between upper- and lower-class students and between students of differing races and ethnic backgrounds. The consensus among researchers is that race and class are two major contributors to the gap and there are several others. Lee and Burkam (2002, ¶ 5) state, "race and ethnicity are closely associated with socioeconomic status." A high proportion of African Americans are categorized as having low socioeconomic status. Therefore, high proportions of African American students generally score lower on standardized tests. Although race and low socioeconomic status do contribute to the gap, Rothman (2001-2002) notes even in suburban schools white students outperform their African American peers.

The National Assessment of Educational Progress (NAEP) data from 1971-1996 shows that the black-white reading gap shrank by almost half and the math gap by a third (Jencks & Phillips, 1998). But in the 1990s the gap for fourth-grade reading and eighth-grade mathematics began to widen again (Haycock, 2001). According to NAEP, white fourth graders scored an average of 30 points higher than their African American peers in 1998. Thirty-eight percent of all whites scored at the proficient level or above, yet only 9% of African Americans scored at this level in reading. A 40-point achievement gap occurred in eighth grade mathematics between whites and African Americans. Thirty-four percent of the nation's white eighth graders scored at the proficient level or above, yet only 5% of their African American peers scored at the same level in mathematics (Education Trust, 2002-2003a, 2002-2003b).

Following the passage of the No Child Left Behind Act (Pub. L. No. 107-110, 2002), many states and districts have made increased efforts to close the achievement gap. Haycock (2001) states that setting standards for what students should learn is key to solving the problem. These standards should set a clear guide for students, parents, teachers, and administrators. Kentucky was one of the first states to adopt a standard-based reform over 12 years ago. This adoption produced the Kentucky Learner Goals and the expectation was that all children would meet these goals. Kentucky officials admit that all students are not meeting these goals yet, but "their progress is clear and compelling" (Haycock, 2001, ¶ 19). In reading, 7 of the 20 top-performing Kentucky schools are high poverty, in math 8 of the 20, and in writing 13 of the top 20 are high poverty (Haycock, 2001).

Related Research

An Indiana school district launched a planning process in the early 1990s to help ensure all students succeed. The process included 600 members of the community: district staff, parents, and students. During a year's time period, the group developed ten strategies to guarantee all students succeed. Two of those strategies included creating a student-centered environment and developing a core curriculum. This school system also provided tutoring during school hours and a reading recovery program for first graders who were not achieving on grade level. With these programs in place, the district's graduation rate has improved by 8% since 1995 and the African American dropout rate has decreased by 7%. Also, the reading, language, and math achievement gaps at different grade levels have closed by 30 to 40% from 1998 to 2001 (Fowler-Finn, 2002).

A magnet middle school in Houston, Texas, earned an "acceptable" rating under the state accountability system. In this school, 98% of white students passed the state tests in 1995-96 and only 82% of African Americans passed. Armed with this information, the school decided to provide additional instruction in mathematics and an after school reading and writing program to help close the achievement gap. After implementing these new programs, 89% percent of African American students in the school passed the state tests in the 2000-01 academic year (Rothman, 2001-2002).

Researchers have noted that achievement gap differences may be due to the kinds of teaching that occurs in classrooms. Waxman and Huang (1997) investigated classrooms in high poverty and high-minority schools. They found that classrooms rated as ineffective/inefficient had 40% of class time where there was no interaction and students were more engaged in whole-class activities. In schools rated effective/efficient, students were interacting with their teachers about 70% of the time and more engaged in individual, student-centered learning.

A study conducted in Canada yielded similar results (Henchey et al., 2001). These researchers found that teachers possessed positive attitudes and high expectations, structured classroom instruction, recognized the need to be accountable for performance, and understood that they had to be innovative if the future of the school was to be assured. There also was a focus on student needs, academic achievement, and other indicators of success. The schools provided assistance and support for students and teachers and also provided a variety and flexibility of structures, programs, and services (Henchey et al., 2001). Knight and Smith (2003) suggest that teachers who are successful in closing the achievement gap exhibit instructional behaviors that involve teaching for meaning as well.

Evaluation of the KY Extended School Services Program

The Extended School Services (ESS) program was established in 1990 as part of the Kentucky Education Reform Act (KERA). Designed specifically to address the needs of Kentucky's at-risk student population, ESS is an aggressive, proactive program for addressing academic problems before they become ingrained (Nesselrodt & Schaffer, 2000a, 2000b). The ESS program extends the school day, week, or year for students at risk of academic failure, providing them with additional instructional time to help them meet academic goals. Rather than

being an "add-on" or "stand-alone" program, ESS is designed to be an integral part of each school's regular academic program, thus ensuring that students receive instructional assistance in core content subjects in which they are performing poorly.

In the fall of 2001, KDE contracted with a partnership of AEL and Western Kentucky University for a comprehensive evaluation of the ESS program during the 2001-02 school year. All learnings from a previous pilot-test evaluation were incorporated into AEL's evaluation design. Fifteen evaluation questions were assembled into five major topics: (1) identification, referral, and assignment of services; (2) profiles of students receiving services; (3) profiles of ESS programs and their implementation patterns; (4) services to students placed at risk; and (5) ESS implementation patterns and outcomes.

AEL's evaluation of the ESS program utilized two major components—statewide surveys and site visits. These components were broken down into five main phases: statewide surveys, training session for site visits, fall/winter site visits, summer visits, and data analyses. The surveys were administered to the district and school ESS coordinators in the fall of 2001. See Figure 1 for a graphic portrayal of AEL's evaluation of the ESS program.

Purpose and Objectives

This study is a secondary analysis of the data from the classroom observations from the fall/winter and summer site visits. The purpose of this study was to explore the classroom observation data from the ESS evaluation site visits in terms of differences across the schools with a minimum achievement gap and those with a large achievement gap. The objectives of this study were to

- dissagregate the classroom observation data from the full ESS evaluation database and split those data by the minimum- and large-gap schools,
- discover any statistically significant differences between the two achievement gap groups for the variables in the classroom observation instruments, and
- compose a picture of the important environmental and instructional differences in minimum and large achievement gap classrooms.

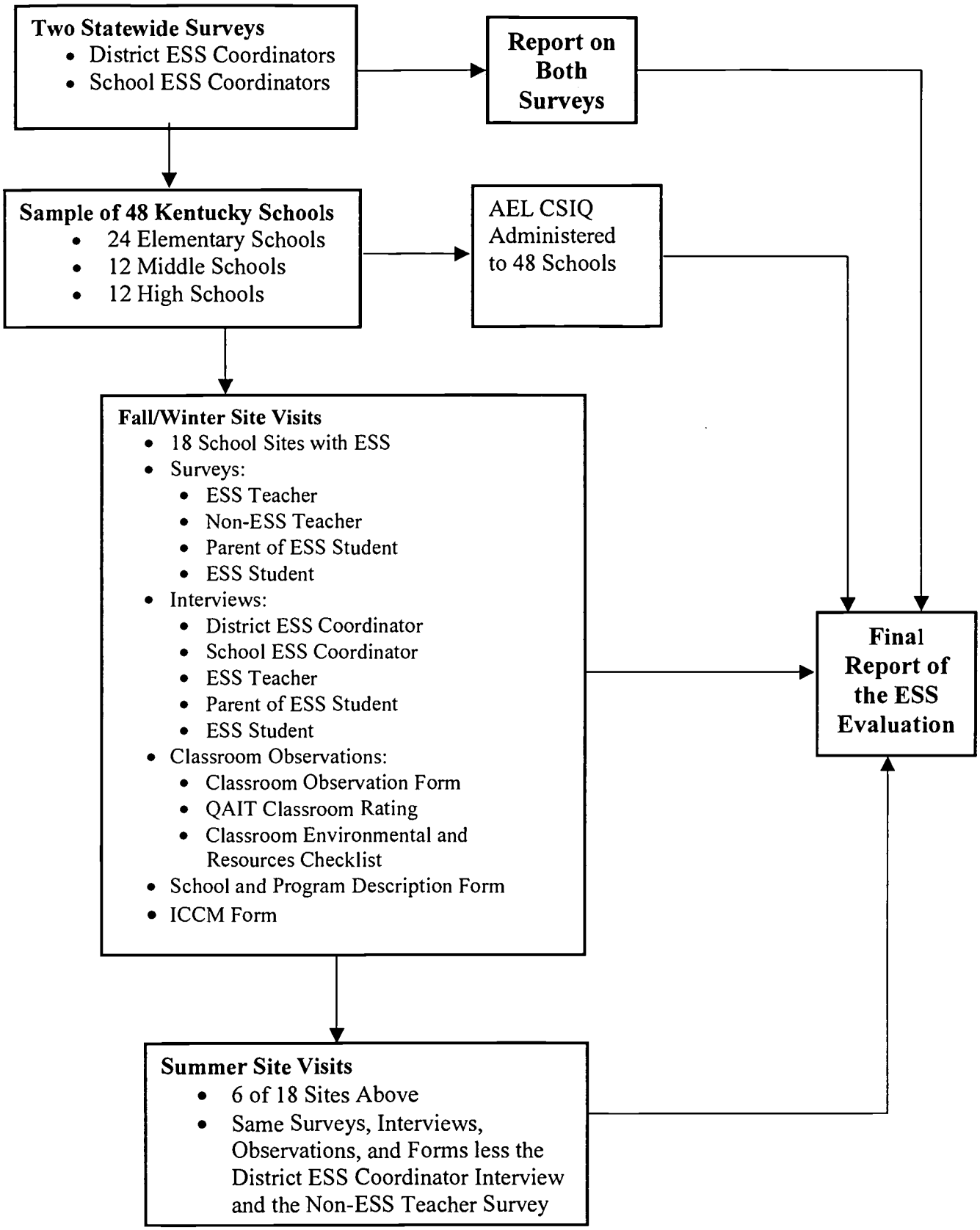

Figure 1: Graphic Portrayal of the Evaluation of the Kentucky Extended School Services Program

METHODS

This sections presents the samples, observation instruments, data collection procedures, and the data analyses processes of this study.

Samples

Schools. A two-stage sampling process was implemented to identify the 24 schools to host the evaluation site visits, of which classroom observations were one data collection process. In the first stage, KDE staff identified a pool of 48 schools through a six-step process that included reviews of student achievement data, percentage of students eligible for free or reduced-price meals, overall academic index score, ethnicity, school-level performance indicators such as novice-level readers and dropout rates, comparisons of subsets of student scores within schools, and geographic and demographic characteristics. This pool of schools included 24 elementary, 12 middle, and 12 high schools.

AEL staff completed the second stage of school sampling by securing the Johnson locale codes (National Center for Education Statistics, 2001) and published enrollment figures (Quality Education Data, 1998) for each of the 48 schools. Using a combination of building level, geographic, urbanicity, and enrollment information, AEL staff selected the 18 schools for the fall/winter 2001-02 site visits. These 18 schools were evenly divided on the "subsets of student scores within schools" (Cowley et. al., 2002). After the 18 fall/winter site visits were completed, AEL staff collaborated by telephone with KDE staff to identify which of the original 48 schools would receive the summer 2002 site visits. It was decided that it would be best for the evaluation to revisit some of the 18 schools selected for the site visits. Based on geography, building level, and general representativeness of the Kentucky ESS program, 6 of the 18 were selected for the summer site visits.

The 48 schools in the original pool were divided into two even groups based on the "comparisons of subsets of student scores within schools" (Cowley et al., 2002), as stated above. In one group were schools where students in ESS or free/reduced-price lunch and African American students were all scoring within 10 points of the school average. Henceforth in this paper, these will be known as the "minimum-gap" schools. Table 1 displays descriptive information on the set of nine minimum-gap schools selected for site visits, including classroom observations. There were three elementary, three middle, and three high schools in this minimum-gap group. School enrollment ranged from 270 to 1,550 students and the free/reduced lunch eligibility of those students ranged from 14% up to 72%. Four of the nine schools were located in rural locales and another three were located in small town locales.

In the second group of schools were students in ESS or free/reduced-price lunch and African American students who were all scoring more than 10 points below the school average. From this point in this paper, these will be known as the "large-gap" schools. Table 2 displays descriptive information on the nine large-gap schools selected for site visits. As above, there were three elementary, three middle, and three high schools in this group. School enrollment

Table 1: Descriptive Information on Minimum-Gap Schools Selected for Site Visits

School Name	Visit Time	Grades [a]	Rounded Enrollment [a]	Attendance Rate [a]	Free/Reduced-Price Lunch (Title I) [b]	Student to Computer Ratio [a]	Locale Type [c]
Minimum Gap - 1	Fall and Summer	9-12	1,330	94.7%	17%	5.2:1	Mid Size City
Minimum Gap - 2	Fall	9-12	1,550	96.0%	14%	4.6:1	Urban Fringe of Large City
Minimum Gap - 3	Fall and Summer	PK-6	270	94.1%	72%	6.0:1	Rural, outside MSA[d]
Minimum Gap - 4	Fall	K-6[1]	300	96.2%	27%	3.3:1	Rural, outside MSA
Minimum Gap - 5	Fall and Summer	9-12[1]	950	96.4%	54%	5.0:1	Small Town
Minimum Gap - 6	Fall	6-8	420	94.6%	50%	3.9:1	Small Town
Minimum Gap - 7	Fall	K-5[1]	390	95.7%	42%	5.1:1	Small Town
Minimum Gap - 8	Fall and Summer	6-8[1]	700	94.5%	18%	4.2:1	Rural, inside MSA
Minimum Gap - 9	Fall	7-8	450	94.1%	38%	5.1:1	Rural, outside MSA

[a] From "School Report Cards," by Kentucky Department of Education, 2002, Retrieved June 2, 2003, from Kentucky Department of Education website: http://aaps.kde.state.ky.us/report_card/

[b & l] From *QED 2001-2002 State School Guides: Kentucky (19th ed.),* by Quality Education Data, 2001, Denver, CO: Quality Education Data.

[c] From Public Elementary/Secondary School Universe Data in *Common Core of Data,* by National Center for Education Statistics, 2002, Washington DC: Author. Available from http://www.nces.ed.gov/ccd/pubschuniv.asp

[d] MSA = Metropolitan Statistical Area

Table 2: Descriptive Information on Large-Gap Schools Selected for Site Visits

School	Visit Time	Grades [a]	Rounded Enrollment [a]	Attendance Rate [a]	Free/Reduced-Price Lunch (Title I) [b]	Student to Computer Ratio [a]	Locale Type [c]
Large Gap - 1	Fall and Summer	9-12	960	95.1%	31%	3.5:1	Large Town
Large Gap - 2	Fall	K-5	470	95.5%	42%	4.7:1	Large City
Large Gap - 3	Fall	3-8	370	94.6%	39%	4.0:1	Small Town
Large Gap - 4	Fall	K-5[1]	640	96.0%	31%	3.9:1	Large Town
Large Gap - 5	Fall	9-12	1,550	92.8%	15%	3.0:1	Mid-size City
Large Gap - 6	Fall	9-12[1]	1,760	89.0%	16%	5.0:1	Mid-size City
Large Gap - 7	Fall	6-8	700	94.2%	16%	5.0:1	Urban Fringe of Large City
Large Gap - 8	Fall and Summer	K-5[1]	540	95.5%	47%	4.0:1	Small Town
Large Gap - 9	Fall	6-8[1]	750	95.7%	15%	6.0:1	Urban Fringe of Large City

[a] From "School Report Cards," by Kentucky Department of Education, 2002, Retrieved June 2, 2003, from Kentucky Department of Education website: http://aaps.kde.state.ky.us/report_card/

[b & 1] From *QED 2001-2002 State School Guides: Kentucky (19th ed.),* by Quality Education Data, 2001, Denver, CO: Quality Education Data.

[c] From Public Elementary/Secondary School Universe Data in *Common Core of Data,* by National Center for Education Statistics, 2002, Washington DC: Author. Available from http://www.nces.ed.gov/ccd/pubschuniv.asp

ranged from 370 to 1,760 students and the free/reduced lunch eligibility of those students ranged from 15% to 47%. There was a variety of locale types but no rural locales were included.

Students/Classes. Students—and their classes—were selected to be observed through a purposeful sampling process. Following the lead from the ESS evaluation pilot test, each school ESS coordinator was asked to select three students to be observed in regular and ESS classes for approximately half of one school day by the data collector. These coordinators were asked to select one student who was making exceptional progress in the program, one student who was making typical progress in the program, and a third student who was making slower than normal progress in ESS. The projected number of students to be observed was 54 in the fall/winter site visits and 18 in the summer visits.

The sampling of classes to be observed was a direct function of schedules of the students (three per site visit) selected by the school ESS coordinators. That is, the purposeful sampling of students to be observed dictated which classes were observed. Observers were instructed to observe the selected students for about half of one day in their regular classes and also their ESS classes (before school, after school, or evening classes). Observers were instructed to follow each targeted student to an English/reading class and a mathematics class for sure. Due to the shadowing scheme for student observations, the precise number of their classes that would be observed was impossible to predict before the observations began, but a low-end estimate was 108 in the fall/winter and 36 in the summer.

In the end, because of the site selection decision for the summer visits, more visits and classroom observations were completed in minimum-gap schools than in large-gap schools. Thirteen site visits were completed in minimum-gap schools and eleven site visits were completed in large-gap schools. The total number of classes with some observation data was 213; however, as is typical in these studies, not all instruments had complete data. For example, specific instruments may not have been completed such as the end-of-session instruments if the student was suddenly called out of the class for a good reason. Thus, there was some missing observation data, but it was not judged to be material. A total of 213 classrooms were observed and formed the database for this study; 193 during the fall/winter and 20 during the summer.

Instruments

Special Strategies Observation System. For the collection of regular classroom and ESS session observation data, three instruments were selected and employed in the pilot test and used in the evaluation. All three were developed and employed in prior research and evaluation studies, refined by AEL and KDE staff, and converted to a scannable format. The three instruments comprise the Special Strategies Observation System (SSOS), which is designed for use in a variety of settings to systematically collect data on essential elements of classroom behavior related to instruction, management, and context. The SSOS is a viable instrument for school effectiveness research due to its strong grounding in the current literature on effective teaching and its utilization of a variety of methodologies. This combination of instruments generates low-, moderate-, and high-inference data; this triangulation of information further

documents the veracity of the data collected. Each instrument that makes up the SSOS is described below.

- *Classroom Observation Form (COF):* The COF is a combination observation system that is best described as a category system with low inference items and includes multiple coding procedures (Nesselrodt & Schaffer, 1993; Sullivan & Meehan, 1983). It is based on the Classroom Activity Record designed by Everston and Burry (1989) and the Stallings Observation System (Stallings, 1980). The top page of the COF collects typical demographic information such as the school, observer, date, number of adults and students in class, subject being observed, and type of class (ESS or regular). The observations occur over 56 minutes, during which the observer switches between coding the entire classroom and focusing on a single student previously selected. Each of seven pages corresponds to eight minutes of class time. The first minute per page—the "snapshot"—looks at student engagement (i.e., the number of students on task, off task, out of the room, or waiting) and grouping strategies (i.e., whether clustered in teacher, aide, or student groups and type of involvement such as working alone, management, interaction, or socialization). The remaining seven minutes per page focus specifically on the target student and include coding one of 27 discrete activities for each minute. See Table 3 for the names of the 27 activities in the four main COF categories.

- *QAIT assessment of classroom:* This instrument is best described as a moderate and high-inference simple coding rating device. QAIT stands for Quality of Instruction, Appropriate Level of Instruction, Incentive, and Use of Time. Fitting on two sheets, it contains 40 items grouped under the four major categories. Each item uses a Likert-type rating scale of 1 to 5 (*Unlike this class* to *Like this class*). This instrument was to be completed at the end of each observation session.

- *Classroom Environment and Resources Checklist:* The Classroom Environment and Resources Checklist (CERC) is a low-inference simple coding sign system. Printed on the front of one sheet, it contains 12 classroom attributes that are coded either as present or not present, such as adequate lighting, use of multi-racial materials, posted assignments, etc. Next, 18 classroom resource items such as textbooks, computers, and worksheets are listed; observers indicate whether such resources are visible or not. If they are, observers indicate whether they are used during the observation. This instrument also was to be completed at the end of each observation session. See Table 4 for the list of items in the CERC.

See Figure 2 for a graphic depiction of the one-hour classroom observation process.

These three observation instruments were tested and utilized in the pilot test by Nesselrodt and Schaffer (2000a, 2000b). Thus, these instruments possess face and content validity and have proven their utility in prior research. A high degree of inter-rater reliability was achieved among the data collectors, given that every participant in the training session passed at or above the 85% criterion of the COF coding assessment held at the conclusion of the training. To assess the degree of internal consistency reliability, Cronbach alpha coefficients

Table 3: Classroom Observation Form Activities by Main Categories

Main Categories	Activities
Teacher-Led	Teacher presentation of content Recitation or discussion Directions for assignments Small-group instruction Tests Checking Praising class
Management/Organization	Procedural or behavioral presentation Administrative routines Transitions Monitoring
Student-Led	Individual seatwork Individual seatwork at computer Pairs or group seatwork Pairs or group seatwork at computer Sustained writing or composition Sustained reading Hands-on learning Independent inquiry or research Student-initiated questions
Off Task	Teacher non-academic activity Waiting time Discipline Student non-academic activity Not occupied Off task Out of room

Table 4: Classroom Environment and Resources Checklist Items

Main Category	Items
Classroom Environment	Use of multi-racial materials
	Use of non-sexist materials
	Posted classroom rules
	Posted assignments
	Cheerful and inviting classroom
	Distinct activity centers
	Adequate lighting
	Comfortable ventilation/temperature
	Student work displayed
	No distracting internal noises/interruptions
	No distracting external noises/interruptions
	Open, risk-free environment
Classroom Resources	Textbooks
	Workbooks/activity books
	Worksheets
	Journals/learning logs
	Classroom library
	Reference materials
	Map and/or globe
	Games and/or puzzles
	Instructional aids/props
	Science/lab table(s)
	Classroom chalkboard
	Student-used equipment
	Overhead projector
	Television
	Computer
	Student manipulatives/hands-on materials
	Audio resources (i.e., tapes, CDs, players)
	Video resources (i.e., tapes, discs, players)

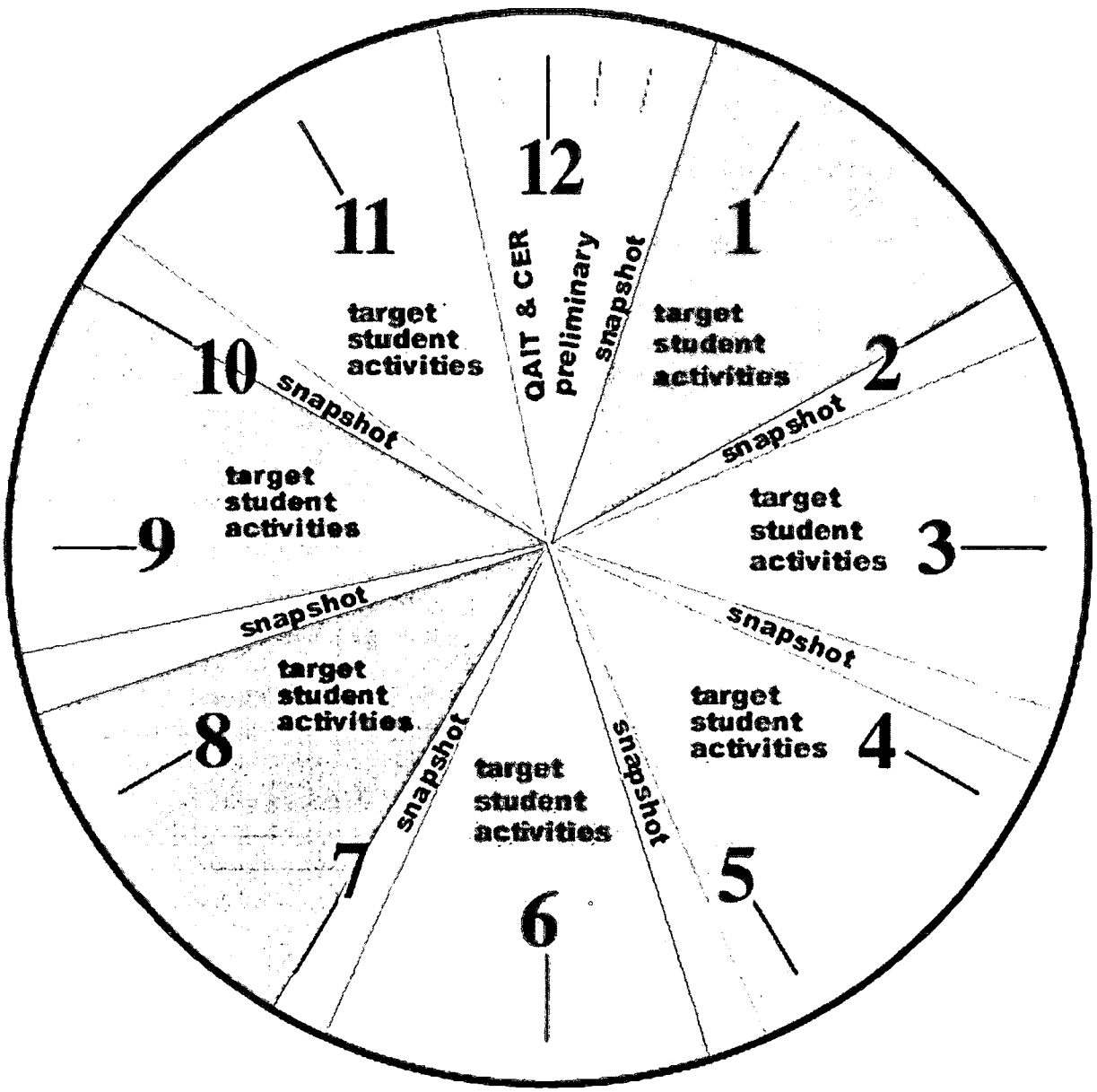

Figure 2: Graphic Depiction of the One-Hour Classroom Observation Process

were computed for this administration of the COF and QAIT instruments, excluding demographic-type items. This procedure was not appropriate for the CERC instrument, given its lack of variance in response options of either selected or not selected. For the COF instrument, this administration of the grouping strategy items resulted in an alpha coefficient of .54; for the student engagement items, a .82. For the QAIT instrument, this administration of all items resulted in a coefficient of .94; by subscale, the coefficients were .91 for quality of instruction, .74 for appropriate level of instruction, .88 for incentive, and .80 for use of time.

Data Collection

Class observations were completed as one component of data collection conducted during the site visits. The 18 fall/winter site visits were completed by the end of February 2002 and the 6 summer visits all were conducted during June 2002.

Procedurally, the pair of data collectors/observers asked the school ESS coordinators to select the three students to be shadowed as one of the first tasks in the site visit. Alternates were sought at this point, anticipating some replacement due to absences, etc. Once the students were selected, their daily schedules for the days the observers were onsite were obtained. With the students' daily schedules as input, the pair of data collectors decided who was going to observe each student and when. The COF segments were completed during the observation period and the QAIT and CERC instruments were completed at the end of the observation period.

At the conclusion of each site visit, data collectors returned all materials, including the observation instruments, to AEL. Each set of returned site visit materials was logged in and checked for completeness in preparation for data analyses.

Data Analyses

After the fall/winter site visits were completed, AEL staff designed data entry templates using Remark scanning software. SSOS data were scanned by school; data files were then cleaned and exported to SPSS for statistical analyses. School files were merged into one master file before analyses began.

COF data were averaged across the number of eight-minute intervals per each observation. Percentages of time for the classroom snapshots and target student activities were calculated for both the minimum- and large-gap groups. Data were analyzed using the 27 individual categories and by grouping these into four main categories: teacher-led, student-led, management/organization, and off-task. Student engagement data and time spent by the target student in the four main instructional categories also were analyzed for the minimum- and large-gap groups. Independent *t*-tests were computed on all these COF variables and, as appropriate, effect sizes were calculated for significant differences. The alpha level was set at .05.

QAIT data were analyzed by the 40 items individually and by the four main subscales composed of their items. Descriptive statistics were used to describe results for both minimum- and large-gap groups. Next, independent *t* tests were conducted to determine if a statistically significant difference existed between the achievement gap groups for the 40 individual items

and each of four main categories: quality of instruction, appropriate level of instruction, incentive, and use of time. Again, alpha was set at .05 and effect sizes were calculated for significant differences.

CERC data were analyzed by calculating crosstab frequency percentages showing whether the classroom attributes were present in the minimum- and large-gap groups. As well, frequency percentages were calculated to show whether various classroom resources were visible and used during the observations. Next, chi-square tests of independence were computed on the classroom environment and resources variables by the minimum- and large-gap groups, and Cramer's V value was used for the effect size.

FINDINGS

This section presents the findings from the classroom observations in the minimum- and large-gap high-performing schools in Kentucky. These findings are presented first by classroom environment and resources, followed by instructional and teaching results. Only statistically significant differences between the two achievement gap groups for each instrument are presented.

Classroom Environment and Resources Results

Table 5 displays the statistical information for the significant chi-squares for classroom environment items by the minimum- or large-gap groups. Of the 12 classroom environment issues, significant differences were found on four items. Those four environmental items were use of multi-racial materials; cheerful and inviting classroom; no distracting external noises/ interruptions; and open, risk-free environment. As shown in the last column in Table 5, in each case, the group with the larger number of times the item was selected than expected was the minimum-gap group. That is, the significant differences for all four classroom environment items favored the minimum-gap group over the large-gap group. However, in terms of practical importance of the significant differences, the Cramer's V values were all .25 or less, or small in Cohen's (1977) suggested qualitative scheme for interpreting effect sizes.

Table 6 displays the statistical information for the significant chi-squares for classroom resource items by the minimum- or large-gap groups. Of the 18 classroom resource items and the dual checkoffs of visible or used during the observation, significant differences were found for three items: textbooks, workbooks/activity books, and overhead projector. For two of the resources (workbooks/activity books and overhead projector), the differences were that they were visible in the classroom while for textbooks, the difference was that they were observed being used. For two of the resources (textbooks and workbooks/activity books), the difference in times observed favored the minimum-gap group; for the overhead projector, the difference favored the large-gap group. As above, the effect sizes all were under .25, or less than small in Cohen's scheme.

QAIT Instrument Results

The QAIT classroom rating results are presented first by the individual items making up the four subscales and second by the subscale level.

Table 7 displays the QAIT rating results for the significant differences of items in the quality of instruction subscale by the minimum- and large-gap groups. Of the 12 rating items in the quality of instruction subscale, significant differences were found on the nine items listed in the left column. All of the items' differences were in favor of the minimum-gap group, which had only one item mean rating under 3.00 on the 5-point rating scale. The differences in the mean ratings ranged from .38 (reminds students of previously learned materials) to .87 (teacher uses an appropriate pace to cover content). The probabilities ranged from .000 up to .048.

Table 5: Statistical Information for Significant Chi-Squares for Classroom Environment Items by Minimum- or Large-Gap Groups

CERC Item	Degrees of Freedom	Chi-Square Value*	Cramer's *V* Value	Group with Larger Number Than Expected
Use of multi-racial materials	1	8.38	.20	Minimum
Cheerful and inviting classroom	1	4.53	.15	Minimum
No distracting external noises/ interruptions	1	6.19	.17	Minimum
Open, risk-free environment	1	12.63	.25	Minimum

* $p < .05$

Table 6: Statistical Information for Significant Chi-Squares for Classroom Resource Items Either Visible or Used by Minimum- or Large-Gap Groups

CERC Item	Visble or Used	Degrees of Freedom	Chi-Square Value*	Cramer's *V* Value	Group with Larger Number Than Expected
Textbooks	Used	1	7.07	.18	Minimum
Workbooks/ activity books	Visible	1	8.31	.20	Minimum
Overhead projector	Visible	1	5.52	.16	Large

*$p < .05$

Table 7: Significant Differences in Mean Ratings in Quality of Instruction Subscale Items by Minimum- and Large-Gap Groups

Q of I Subscale Item	Minimum Gap		Large Gap		*df*	*t*	*p*	Diff.	Cohen's *d*
	Mean	Std. Dev.	Mean	Std. Dev.					
Organizes information in an orderly way	4.03	1.28	3.45	1.40	203	3.07	.002	.58	.43
Notes transitions to new topics	3.43	1.52	2.80	1.53	202	2.95	.004	.63	.41
Frequently restates essential principles	3.78	1.49	3.17	1.52	202	2.87	.005	.61	.40
Uses devices such as advanced organizers	2.85	1.72	2.24	1.40	202	2.75	.006	.60	.39
Reminds students of previously learned materials	3.77	1.41	3.39	1.31	204	1.99	.048	.38	.28
The teacher shows a sense of humor	3.38	1.45	2.83	1.51	206	2.67	.008	.55	.37
Conducts formal and/or informal assessments	3.95	1.27	3.19	1.52	204	3.89	.000	.76	.54
Provides immediate and corrective feedback	4.10	1.23	3.48	1.49	205	3.27	.001	.62	.46
Teacher uses an appropriate pace to cover content	4.26	1.03	3.39	1.47	205	4.88	.000	.87	.70

Effect sizes, as measured by Cohen's *d* value (1988) ranged from .28 to .70. In terms of Cohen's qualitative scheme for these effect sizes, all were somewhere between small and large (two were medium), indicating that all of them did possess some practical importance.

Table 8 presents the QAIT rating results for the significant differences of items in both the appropriate level of instruction and incentive subscales. Of the seven rating items in the appropriate level of instruction subscale, only one significant difference was discovered and it was for the "Uses individualized instruction" item. With a mean rating of 2.69 on the 5-point scale, the minimum-gap group was .51 points higher than the large-gap group. The probability was .023 and the Cohen's *d* value was .32, which is small in qualitative terms. Of the 18 items in the incentive subscale, significant differences were found on three items, all favoring the minimum-gap group. The mean scores for two of those items for the minimum-gap group were above 3.0 (communicating high expectations and efforts by the students lead to success) and one was under 3.0 (relating topics to students lives). The differences in the means ranged from .42 to .57 and the probabilities ranged from .008 to .011 (tie). The three Cohen's *d* values were similar at .35, .36, and .37, respectively, indicating some practical importance.

Table 9 shows the QAIT observation rating results for the use of time subscale items. This subscale consists of three items and significant differences were found on all three items in favor of the minimum-gap group. Although the mean ratings for both groups were higher than the items on the other subscales, the significant differences in those ratings were .33 (students attend to lessons), .47 (necessary time is allocated for instruction), and .48 (the teacher uses effective management). The probabilities ranged from .005 to .024. The Cohen's *d* values were .32, .39, and .40, indicating that the differences had some practical importance.

Another way to present the QAIT observation results is at the subscale level with all the items' ratings averaged. Table 10 displays the QAIT subscale results. Two subscales (quality of instruction and use of time) yielded significant differences, both favoring the minimum-gap group. The difference in mean scores was .51 for quality of instruction and .43 for use of time. The probabilities were .51 and .43, respectively, and their Cohen's *d* values were similar at .51 and .45. The qualitative descriptor for the quality of instruction *d* value is medium and is small for use of time. These effect sizes indicate that there was some real, practical importance to the statistically significant differences on these two subscales.

Classroom Observation Form Results

Results from the COF observations are presented in terms of the number of minutes. There were 27 individual categories of activities, four main categories of those 27 activities collapsed, four major groupings of subjects in the classrooms, and four categories of off-task/on-task behaviors coded. The 27 individual activities and four collapsed categories were coded in terms of what the selected students were doing, while the latter 8 categories were snapshots of the full classroom. Additionally, the total number of minutes observed for the students was computed. Table 11 displays the significant differences in observed minutes of instruction by the minimum- and large-gap groups. Of the 40 possible COF categories, significant differences across groups were found on just three of them (administrative routines, total number of

Table 8: Significant Differences in Mean Ratings in Appropriate Level of Instruction and Incentive Subscale Items by Minimum- and Large-Gap Groups

A.L. of I. and Incentive Subscale Items	Minimum Gap		Large Gap		*df*	*t*	*p*	Diff.	Cohen's *d*
	Mean	Std. Dev.	Mean	Std. Dev.					
Appropriate Level of Instruction Subscale									
Uses individualized instruction	2.69	1.64	2.18	1.55	207	2.30	.023	.51	.32
Incentive Subscale									
Relating topics to students lives	2.70	1.59	2.14	1.47	206	2.67	.008	.57	.37
Communicating high expectations	3.34	1.46	2.82	1.49	207	2.57	.011	.52	.35
Efforts by the students lead to success	3.89	1.09	3.47	1.27	205	2.56	.011	.42	.36

Table 9: Significant Differences in Mean Ratings in Use of Time Subscale Items by Minimum- and Large-Gap Groups

Use of Time Subscale Items	Minimum Gap		Large Gap		*df*	*t*	*p*	Diff.	Cohen's *d*
	Mean	Std. Dev.	Mean	Std. Dev.					
Necessary time is allocated for instruction	4.18	1.08	3.81	1.32	205	2.77	.006	.47	.39
The teacher uses effective management	4.30	1.02	3.82	1.37	205	2.84	.005	.48	.40
Students attend to lessons	4.26	0.90	3.93	1.20	205	2.27	.024	.33	.32

Table 10: Differences in Mean Ratings on Four Subscales of the QAIT Instrument

by Minimum- and Large-Gap Groups

QAIT Subscale Name	Minimum Gap		Large Gap		*df*	*t*	*p*	Diff.	Cohen's *d*
	Mean	Std. Dev.	Mean	Std. Dev.					
Quality of Instruction	3.64	0.96	3.13	1.03	208	3.69	.000	.51	.51
Appropriate Level of Instruction	2.44	1.02	2.34	0.93	207	0.74	.460	.10	.10
Incentive	2.43	0.71	2.27	0.75	208	1.63	.105	.17	.23
Use of Time	4.28	0.75	3.85	1.14	205	3.15	.002	.43	.45

Table 11: Significant Differences in Observed Minutes of Instruction by Minimum- and Large-Gap Groups

Classroom Observation Form Category	Minimum Gap		Large Gap		*df*	*t*	*p*	Diff.	Cohen's *d*
	Mean	Std. Dev.	Mean	Std. Dev.					
Administrative routines	1.96	1.52	3.63	2.34	41	-2.84	.007	-1.67	-.87
Total number of observed minutes	47.14	8.71	42.73	10.16	211	3.38	.001	4.41	.47
Student-led activities	22.06	13.44	18.06	12.27	164	2.01	.046	4.00	.31

observed minutes, and collapsed student-led activities). In each case, the difference favored the minimum-gap group. For example, the large-gap group was observed spending an average of 1.67 minutes more on administrative routines than the minimum-gap group. This difference in minutes was significant at the .007 level. Further, Cohen's *d* value for this difference was -.87, large in Cohen's (1988) scheme, indicating much practical importance to that difference. With respect to the total number of minutes observed, the difference was 4.41 minutes more for the minimum-gap group and the probability of that difference was .001. Cohen's *d* value for total minutes was .47, quite nearly at the medium level. Thus, the difference in observed minutes was rather important in a practical sense, in addition to the statistical sense. Finally, there was a significant difference in the number of minutes of student-led activities. There was a four-minute difference in student-led activities in favor of the minimum-gap group for the 164 classes where this code was observed. The probability of obtaining this difference was .046 and the Cohen's *d* value was .31, in the small category. Thus, we can say that the difference in minutes of student-led activities in favor of the minimum-gap group had some practical importance in addition to its statistical significance.

CONCLUSIONS AND RECOMMENDATIONS

Based on the findings from this study, certain conclusions and recommendations are warranted. First, it should be noted that the study was of the descriptive information differences between minimum and large achievement gap schools based on systematic observations of classrooms in those schools. Thus, causal inferences are not being claimed. Nonetheless, the observed differences are real. This section is organized by several main topics.

Classroom Time

The results of this study show that there were important differences in the use of time in the classrooms of minimum- and large-gap schools. Classroom time was used more efficiently and effectively in the minimum-gap schools. Teachers in the large-gap schools spent more time on administrative routines in their classrooms. Teachers in minimum-gap schools were observed providing instruction to students for more minutes than in large-gap schools. Too, we learned that minimum-gap teachers deployed more time for student-led activities than their counterparts. Teachers in minimum-gap schools were observed using effective management techniques and allocating necessary time for instruction. Also, students in those classes were observed attending to those lessons better than their counterparts. Finally, the minimum-gap teachers used an appropriate pace in their classrooms to cover the content.

Although all the schools in this study were identified as high performing in terms of overall academic index scores, the differences in time use in their classrooms is important. The minimum-gap schools should know this as feedback to one of the areas in which they are doing well, and work to continue to do well. But more importantly, the large-gap schools should learn of these results because, on the basis of the overall academic index score, they may feel they really are high-performing schools in all areas. In reality, when compared to similar schools also identified as being high performing but also with subsets of students, the large-gap schools could do much better with the use of their classroom time. Although the actual number of minutes difference between the two groups of schools appears small, when multiplied by the number of hours per day, days per week, weeks per month, and months per school year, the difference in minutes is huge. For example, in terms of classroom time spent on administrative routines, observed teachers in the large-gap schools, on average, spent 24.35 hours more over one school year (1.67 x 5 periods = 8.35/day x 175 days = 1,461 minutes). In the case of student-led activities, observed teachers in the minimum-gap schools spent, on average, 58.33 more hours over the school year.

In addition to informing schools about the differences in the use of classroom time, one recommendation would be to disseminate information of efficient use of time to educators in these schools. For example, there are research-based professional development programs to improve teachers' classroom management and organization skills in order to improve their use of instructional time. These programs, or even demonstrations of some efficient use of time techniques, could be demonstrated to school staff.

Instructional Resources

While the minimum- and large-gap classrooms were observed in regard to the presence and/or use of a wide variety of resources, only a few significant differences were found in this investigation. Textbooks and the use of multi-racial materials were discovered more in the minimum-gap schools. And, while the use of textbooks is understandable, the use of multi-racial materials is a major discovery of this study. One would expect that all high-performing schools had similar uses of multi-racial materials, but this study revealed more use of such materials in the minimum-gap schools. Unfortunately, the observation system did not pinpoint specifically what type of multi-racial materials were used in the classrooms, but their use was noted. Expansion of the use of multi-racial materials seems to be an important recommendation emanating from this study.

The other differences in classroom resources are interesting, but puzzling. Here we are referring to the differences in the two items that were "visible" in the minimum- and large-gap schools: workbooks/activity books and overhead projector. The former favored the minimum-gap schools while the latter favored the large-gap schools. As noted above, we found that minimum-gap school classrooms used textbooks more than their counterparts, but they had more workbooks/activity books visible—not used—than their counterparts. So, one inference is that merely having workbooks/activity books is not as important as actually using textbooks in classrooms.

Classroom Climate

Another one of the interesting findings of this study is the "picture" of the classroom differences between the minimum- and large-gap schools. Here we are talking about the climate or tone of the classrooms that emerged from the observations. The climate in the minimum-gap school classrooms was more cheerful and inviting, open and risk-free, and had less distracting external noises/interruptions. Also, minimum-gap school classroom teachers showed a sense of humor more than their counterparts in large-gap schools. Certainly, taken together, these variables paint a very positive and inviting picture of the minimum-gap classrooms. The inference drawn from this picture of these classrooms is that it is worth the effort to seek to attain and maintain such a climate or tone for the classrooms. The implications for doing so may be more difficult to achieve, however. The recommendation here is transmit this finding to those in charge of supervising classroom teachers in the hopes that much of it can be advocated in classrooms. We are not sure that teaching educators how to have a sense of humor in their teaching is viable as a professional development target. However, we do think that professional development sessions can, and probably should, be on the topics of creating cheerful and inviting classrooms and maintaining open, risk-free environments. Too, the negative effects of distractions and interruptions to classrooms has been well-documented and researched (e.g., Sullivan & Meehan, 1983).

Expectations and Feedback

Clearly, the roles of expectations and feedback to students differed across the two achievement gap groups. Teachers in the minimum-gap schools communicated high expectations to their students, conducted formal and/or informal assessments of their students, and also provided immediate and corrective feedback to students. These actions were significantly less evident in the large-gap schools. These are basic teaching behaviors that have been part of the effective instructional practices scene for many years. It is rather disappointing to discover these particular differences, especially in schools that have been identified as high performing. A natural recommendation is that professional development be provided to staff in the large-gap schools on the topics of high expectations for all students and providing assessments and feedback to students. We know that there has been much progress in Kentucky schools on these topics since the enactment of KERA, but this study showed that there is a way to go yet.

Quality of Instruction

One of the most surprising results of this study was differences in the quality of instruction between the minimum- and large-gap schools. Not only was the quality of instruction subscale itself different in a meaningful way between the two achievement gap groups, but so too were many of the items in the subscale. This was rather surprising and also important. The principles of effective instruction have been well known for many years and the expectation that they have been integral components of teacher preparation programs is not unwarranted. To discover differences on the basic instructional principles is disheartening on the one hand, but illuminating on the other. The illumination resides in that these instructional techniques can be improved. There are models, there is research, there are programs to help teachers learn to use these principles and skills in their teaching.

The recommendation is that some consideration be given to professional development sessions for teachers in schools with large achievement gaps on many of the basic principles of effective instruction. Such topics include organizing information orderly, transitions to new topics, restating essential principles, advanced organizers, and reminding students of previously-learned materials.

Level of Instruction

Finally, this study uncovered differences in the level of instruction between the minimum- and large-gap schools. Although not as far-reaching or pervasive as the differences in the quality of instruction area, there were some important differences between the two groups. In the minimum-gap schools, the classrooms were observed relating topics to students' lives more than their counterparts. Also, it was observed that the efforts by the students led to success. Again, these are crucial instructional techniques that have been known for years. Obviously, some teachers in some schools do better at it than others. This needs to be communicated and discussed among school staff. The recommendation is that the differences on these topics be reduced in the schools.

Overall

In sum, this study provides evidence of the disparity that exists even between schools that are classified as high performing. While the schools in this study were performing at a high level overall, further investigation did illuminate the subtle classroom differences between those schools serving all students well and those that were not as successful with specific subgroups of students.

REFERENCES

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Academic Press: New York.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates: Hillsdale, New Jersey.

D'Amico, J. J. (2001). A closer look at the minority achievement gap. Retrieved June 25, 2002, from www.ers.org/spectrum/spg01a.htm

Education Trust. (Winter 2002-2003a). Education watch Kentucky key education facts and figures. Retrieved June 16, 2003, from www.edtrust.org

Education Trust. (Winter 2002-2003b). Education watch the nation key education facts and figures. Retrieved June 16, 2003, from www.edtrust.org

Everston, C., & Burry, J. (1989). Capturing classroom context: The observation system as lens for assessment. *Journal of Personnel Evaluation in Education, 2*, 297-320.

Fowler-Finn, T. (2002). I think I can. *American School Board, 189,* 36-39.

Haycock, K. (2001). Closing the achievement gap. *Educational Leadership, 58 (6).* Retrieved June 5, 2003, from http://www.ascd.org/readingroom/edlead/0103/haycock.htm

Henchey, N., Dunnigan, M., Gardner, A., Lessard, C., Muhtadi, N., Raham, H., & Violato, C. (2001, November). *Schools that make a difference: Final report* (SAEE Research Series #6-D). Kelwona, BC: Society For The Advancement of Excellence in Education.

Jackson, M. P. (2001). All children can learn: So why is there a minority achievement gap? *Perspectives, 12* (2), 1 & 4.

Jencks, C., & Phillips, M. (1998). America's next achievement gap: Closing the black-white test score gap. *The American Prospect, 9*(40). Retrieved June 6, 2003, from http://www.prospect.org/print/V9/40/jencks-c.html

Kentucky Department of Education. (2002). *School Report Cards.* Retrieved June 2, 2003, from http://aaps.kde.state.ky.us/report_card/

Knight, S., & Smith, R. (2003, April). *Observing teaching for meaning in classrooms of teachers successful in narrowing the achievement gaps.* Paper presented at the meeting of the American Education Research Association, Chicago, IL.

Lee, V., & Burkam, D. (2002). Inequality at the starting gate: Social background differences in achievement as children begin school. Retrieved June 11, 2003, from http://www.epinet.org/content.cfm/books_starting_gate

National Center for Education Statistics. (2001, August). *Common core of data* [Internet file]. Available from National Center for Education Statistics Website, http://nces.ed.gov/ccdweb/school/index.asp

National Center for Education Statistics. (2002). Public Elementary/Secondary School Universe Survey Data in *Common Core of Data, 2000-2001[Data File].* Washington, DC: Author. Available from http://www.nces.ed.gov/ccd/pubschuniv.asp

Nesselrodt, P. S., & Schaffer, E. C. (1993, Atlanta). *The ISERP programme: A revised classroom observation instrument.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Nesselrodt, P. S., & Schaffer, E. C. (2000a). *External evaluation of Kentucky's Extended School Services*: Spring, 2000: Phase I–Final report. Authors

Nesselrodt, P. S., & Schaffer, E. C. (2000b). *External evaluation of Kentucky's Extended School Services*: Spring, 2000: Phase I–Final report-Part 2. Authors.

No Child Left Behind Act of 2001, Pub. L. No. 107-110. (2002).

Quality Education Data. (1998). *Quality Education Data's state school guides: Kentucky 1998 & 1999* (16th ed.). Denver, CO: Author.

Quality Education Data. (2001). *Quality education data 2001-2002 state school guides: Kentucky* (19th ed.). Denver, CO: Author.

Reynolds, G. M. (2002). Identifying and eliminating the achievement gaps: A research-based approach. *Viewpoints, 9*, 3-11.

Rothman, R. (Winter 2001/2002). Closing the achievement gap: How schools are making it happen. *The Jouranl of the Annenberg Challenge, 5 (2).* Retrieved June 11, 2003, from www.annenbergchallenge.org/pubs/cj/gap_cj.htm.

Singham, M. (2003). The achievement gap: Myths and reality. *Phi Delta Kappan, 84*(8), 586-591. Retrieved June 10, 2003, from Ebscohost database.

Stallings, J. (1980). Allocated learning time revisited: Or beyond time on task. *Educational Researcher*, *9*(11),11-16.

Sullivan, D. K., & Meehan, M. L. (1983, April). *The development, pilot-test, and field-test of an instrument for the systematic observation of classroom interruptions.* Paper presented at

the annual meeting of the American Educational Research Association, Montreal, Canada.

Viadero, D. (March 22, 2000). Lags in minority achievement defy traditionally explanations. *Education Week.* Retrieved June 16, 2003, from http://www.edweek.org/ew/ewstrory.cim?slug=28causes.h19

Waxman, H., & Huang, S. L. (1997). Classroom process differences in inner-city elementary schools. *Journal of Educational Research, 91 (1),* 49-60. Retrieved June 12, 2003, from Ebscohost Database.

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

NOTICE

Reproduction Basis

X	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").