



2004-2005 Pilot Test of Two New Culturally Responsive Curriculum Units

Christina L. Gilchrist, Ph.D.
Georgia K. Hughes, M.A.
Joseph L. Holloway, B.S.

December 2005

Appalachia Educational Laboratory (AEL)
at

EDVANTIA™
Partners in education. Focused on results.

2004-2005 Pilot Test of Two New Culturally Responsive Curriculum Units

By

Christina L. Gilchrist, Ph.D.

Georgia K. Hughes, M.A.

Joseph L. Holloway, B.S.

December 2005

Appalachia Educational Laboratory at Edvantia, Inc.
Charleston, West Virginia

Edvantia was founded in 1966 as the Appalachia Educational Laboratory, Inc. (AEL); on September 1, 2005, AEL became Edvantia, Inc. The Regional Educational Laboratory for the Appalachian region is known as the Appalachia Educational Laboratory at Edvantia.

Edvantia is a nonprofit education research and development corporation, founded in 1966, that partners with practitioners, education agencies, publishers, and service providers to improve learning and advance student success. Edvantia provides clients with a range of services, including research, evaluation, professional development, and consulting.

For information about Edvantia research, products, or services, contact



P.O. Box 1348, Charleston, WV 25325 • 304.347.0400 • 800.624.9120 • fax 304.347.0487
One Vantage Way, Suite D-210, Nashville, TN 37228 • 615.565.0101 • fax 615.565.0112
info@edvantia.org • www.edvantia.org

© 2005 by Edvantia, Inc.

All rights reserved. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

Edvantia maintains policies and procedures that govern the Institutional Review Board in ensuring the welfare and protecting the rights of human subjects who choose to participate in research and evaluation activities.

This publication is based on work sponsored wholly or in part by the Institute of Education Sciences (IES), U.S. Department of Education, under contract number ED-01-CO-0016. Its contents do not necessarily reflect the positions or policies of IES, the Department, or any other agency of the U.S. government.

Edvantia is an equal employment opportunity/affirmative action employer.

CONTENTS

EXECUTIVE SUMMARY	iv
INTRODUCTION	1
Background	1
Purpose.....	2
Audience	2
Limitations	3
Overview	3
FOCUS.....	4
Culturally Responsive Curriculum Units.....	4
Fourth Grade: Organisms, Organisms Everywhere!.....	5
Ninth Grade: Unraveling the Details of the Past	6
Pilot Test Evaluation Questions.....	6
METHODS	7
Perspective	7
Design	7
Time Frame.....	7
Sampling.....	7
Data Collection	8
Classroom Observations	8
Survey	8
Unit Feedback	9
Data Analysis	9
FINDINGS	10
Fourth Grade	10
Classroom Observations	10
Survey	11
Unit Feedback	12
Ninth Grade.....	12
Classroom Observations	12
Survey	14
Unit Feedback	14
DISCUSSION.....	15
Lessons Learned.....	15
Participation	15
Unit Quality and Workability	15
Recommendations.....	16

REFERENCES	17
APPENDIXES	19
A: Data Collection Instrument Descriptions	
B: Informed Consent Procedures	

EXECUTIVE SUMMARY

The purpose of this report is to document the pilot testing of two culturally responsive curriculum units (CRCUs) and inform AEL/Edvantia staff about how the units could be improved.

Culturally responsive curriculum units are based on five principles of culturally responsive teaching: high expectations, cultural competence, active teaching, student-controlled discourse, and relevant curriculum instructional practices. One fourth-grade and one ninth-grade unit were developed by Lab staff and Minority Fellows (who were hired to assist with unit development) to complement previously developed and piloted curriculum units.

The county superintendent designated two pilot and, in consultation with lab staff, two matching schools. Fourth- and ninth-grade pilot teachers agreed to pilot a unit during the 2004–2005 school year; matching teachers continued with their original lesson plans. An intact-group comparison of pilot and matching teachers was implemented.

The goals of this pilot test included (1) examining whether the instructional behaviors differed between pilot and matching teachers; (2) investigating whether the perceptions of their schools' capacity for improvement differed between pilot and matching teachers; and (3) soliciting feedback from pilot teachers regarding the workability and quality of the CRCUs that they implemented. Four main data collection activities were planned—classroom observations using the Special Strategies Observation System-Revised (SSOS-R), the Measure of School Capacity Improvement (MSCI) survey, the Curriculum Unit Evaluation Rubric (CUER), and a follow-up teacher interview.

The fourth-grade classroom observation data showed that pilot teachers provided students with incentives for learning and a positive learning environment; the data also favored pilot schools along many constructs. For example, pilot teachers tended to use methods associated with active teaching, while matching teachers tended to use methods associated with direct teaching. Furthermore, pilot teachers tended to have more positive classroom environments, have a higher percentage of visible resources, and use those resources more often than matching teachers. Pilot teachers rated the capacity of their schools for improvement positively; they also rated the majority of items for the unit they taught as *exemplary*, with only one eighth of the items being rated *acceptable* and none being rated *improvement recommended*.

Ninth-grade teachers were observed but did not participate in the other data collection activities, although opportunities were provided for them to give feedback regarding the quality and workability of the units they taught. Classroom observation data indicated that pilot teachers were rated higher than matching teachers on quality of instruction, appropriate level of instruction, incentives for learning, and use of time. Pilot teachers also provided more challenging activities, while matching teachers tended toward using a lower level of instructional conversation. Compared to matching teachers, ninth-grade pilot teachers also had and used more environmental and instructional resources; matching teachers did not use any of the visible resources.

Although sample sizes were small, useful lessons learned and general indicators of unit quality were gleaned. Because pilot test data, by design, are not intended for generalization, the results of this study fit its original purpose—to gather information regarding the quality and workability of units within the intended setting of use.

One lesson learned was that volunteer pilot teachers' schedules for unit implementation and data collection may not agree with that of the researchers. Pilot teachers may not have had any incentive to implement the unit or participate in data collection activities. When they did implement a unit, there was poor coordination among the teachers and the researchers. Such scheduling problems led to the inability to follow up with additional data collection activities.

The researchers also learned about the quality of the CRCUs being tested. Fourth-grade teachers rated the CRCUs that they implemented positively; ninth-grade teachers did not provide any feedback. Classroom observation data suggest that pilot teachers exhibited instructional behaviors more consistent with the principles of culturally responsive instruction than did matching teachers. This observation is further supported by teacher feedback indicating that they thought the unit was responsive to students' cultures. This study investigated one facet of cultural responsiveness; it did not, however, examine students' assessment of cultural responsiveness.

Based on these lessons learned, the researchers suggest the following: (1) conduct a field test, which is the next step in product development research; (2) solicit constructive feedback, even when positive or high ratings are given; (3) obtain written agreements with teachers regarding implementation and data collection activities; (4) in future research studies, to rule out self-selection influences, attempt to randomly select and assign teachers; and (5) continue investigating the extent to which teachers' perceptions of their schools' capacity to improve increases as they continue to implement culturally responsive curriculum units.

INTRODUCTION

This report describes the pilot test that was conducted for two new culturally responsive curriculum units during the 2004-2005 school year.

Background

The No Child Left Behind Act of 2001 requires schools to demonstrate that all students, including economically disadvantaged and minority students, are achieving academic benchmarks. However, indicators of academic achievement repeatedly show that African American students are not faring well (The College Board, 1999). The achievement gap across the nation between African American and White youth began to narrow in the 1970s but has widened in the past decade (D'Amico, 2001; Haycock, 2001; Johnston & Viadero, 2000).

West Virginia, like other states across the nation, is not exempt from the challenges associated with improving the achievement of African American children and youth, many of whom are economically disadvantaged. Although West Virginia's African American student population is small, consisting of approximately 4% of the state's 3rd- through 11th-grade population, a disproportionate percentage of African American students score in the lower quartiles on standardized tests. According to West Virginia's 2001-2002 No Child Left Behind Report Card (West Virginia Department of Education, n.d.), only 39% of the state's African American students in Grades 3-11 scored above the 50th percentile in basic skills on the Stanford 9 Achievement Test. African American students often are overrepresented in special education categories and underrepresented in Advanced Placement courses. Several statewide, regional, and county advocacy groups have campaigned for programs to ameliorate the situation.

The Appalachia Educational Laboratory at Edvantia has been collaborating with the Kanawha County School (KCS) District since 2001 to develop a process for systematically addressing these issues at the district level. In 2001, KCS district officials asked a Lab staff member for help with planning to improve the academic achievement of African American students in the county. That initial request ultimately resulted in a pilot project in which AEL/Edvantia worked with four schools and with district leaders to improve instruction for all students, particularly those who are African American and economically disadvantaged. That project, called the Pilot Schools Project, was a multiyear initiative that culminated in a year-long research project (see Hughes et al., 2004). A brief description of some Pilot Schools Project professional development activities follows.

In May 2001, the KCS management team identified four schools that had high concentrations of African American and economically disadvantaged students and showed significant achievement gaps between racial and SES groups on the Stanford 9 Achievement Test. Team members from all four schools attended professional development sessions to extend learning and to focus on how best to implement culturally responsive schooling and instructional practices in their schools.

During the 2001-2002 and 2002-2003 school years, team members from all four pilot schools attended professional development sessions led by Lab staff and, in some cases, KCS curriculum specialists. These sessions introduced topics such as culturally responsive instruction, peer observation, and reflective teaching practice. A Lab staff member or consultant worked intensely with each of the pilot schools, facilitating the twice-monthly team meetings that were designed to extend learning from the professional development sessions and focus on how best to implement culturally responsive schooling and instructional practices in the school. Lab staff also provided additional assistance in their respectively assigned schools as requested by the principals or teachers.

During the 2003-2004 school year, Lab staff conducted two workshops for Maximizing the Achievement of African American Children in Kanawha (MAACK) teams of school and community members from all four of the Pilot Schools Project pilot schools. The first workshop introduced nine principles of culturally responsive instruction (as defined by the Knowledge Loom, a product of the Education Alliance at Brown University, <http://knowledgeloom.org>) and a template for planning culturally responsive lessons. The second workshop was used to facilitate participants' reflections on their learning about culturally responsive teaching and to gain insights about the levels of implementation of the principles across the four schools. In each pilot school, a team of teachers selected by the principal also taught a culturally responsive curriculum unit developed by Lab staff and Minority Fellows. The units were pilot tested by teachers in Grades K-2, 6, and 11. Using rubric ratings and the information gathered from teachers during debriefings, Lab staff and Minority Fellows revised the three existing units in the summer of 2004.

KCS agreed to implement the units developed for the Pilot Schools Project. The superintendent has stated that he intends to continue taking proactive steps to address the achievement gap; he hopes to have culturally responsive units for every grade level. Knowing the level of interest in culturally responsive instruction, AEL/Edvantia requested in August 2004 that KCS assist with pilot testing of *new* culturally responsive units not tested during the Pilot Schools Project. This report documents the ensuing activities.

Purpose

The purpose of this project was to pilot test two culturally responsive curriculum units not tested during the Pilot Schools Project.

Audience

The primary audience for this report is the Lab staff who designed the curriculum units. A secondary audience consists of the professional staffs from within Kanawha County Schools who would likely be using the units. Finally, the Institute of Education Sciences at the U.S. Department of Education will be given a copy of this report, pursuant to fulfilling the requirements of Edvantia's Regional Educational Laboratory (REL) contract.

Limitations

This study was intentionally limited to the pilot testing of curriculum materials, which typically occurs fairly early in the curriculum development process so that preliminary results can inform subsequent product development and research (Contract Research Corporation, 1975). Therefore, these results are primarily appropriate for internal product development purposes but not for generalization.

Overview

This report describes the (1) culturally responsive curriculum units, (2) methods used in the pilot test, (3) findings, and (4) discussion of lessons learned and recommendations.

FOCUS

This section describes the culturally responsive curriculum units and the evaluation questions that guided the pilot test.

Culturally Responsive Curriculum Units

Culturally responsive teaching is based on the idea that culture is central to student learning. According to Ladson-Billings (1994), “it is an approach that empowers students intellectually, socially, emotionally, and politically by using cultural referents to impart knowledge, skills, and attitudes” (p. 18). Gay (2002) concurs that culturally responsive teaching uses “the cultural characteristics, experiences, and perspectives of ethnically diverse students as conduits for teaching them more effectively” (p. 106). Gay’s sociocultural approach to teaching, based on the work of Russian psychologist Lev Vygotsky, provides instructional scaffolding that encourages students to learn by building on the experiences, knowledge, and skills they bring to the classroom. In order to do this effectively, teachers need to acquire knowledge about the cultural particularities of the groups within their classroom and to transform that information into effective classroom practice (McIntyre, Roseberry, & Gonzalez, 2001). Furthermore, studies of culturally responsive practice reinforce the work of Ladson-Billings, Gay and others. The nine principles of culturally responsive instruction (as defined by the Knowledge Loom, a product of the Education Alliance at Brown University, <http://knowledgeloom.org>) make explicit the process of culturally responsive teaching.

During the 2003-2004 school year, AEL/Edvantia used the Knowledge Loom’s nine principles of culturally responsive instruction as the curriculum for bimonthly meetings with the Pilot Schools Project teams in each of the four pilot schools. The teachers on these teams designed and taught lessons that incorporated the principles and reflected on the lessons’ impact on their students’ learning. Some teachers also taught a culturally responsive curriculum unit developed by Lab staff and Minority Fellows.

At the end of the 2003-2004 school year, learnings were gleaned regarding the teams’ understanding and application of the nine principles in their teaching. Analyses of these learnings led AEL/Edvantia staff to further distill the principles of culturally responsive instruction into five principles:

1. **High expectations.** School staff consistently communicate that they believe in students’ ability to succeed and to achieve mastery of challenging standards of learning.
2. **Cultural competence.** Educators value students’ cultures, beliefs, and families and incorporate them into instructional and schooling practices.
3. **Active teaching.** Teachers facilitate learning by engaging students in a variety of learning activities.

4. **Student-controlled discourse.** Teachers create classrooms that invite personal dialogue—between students and between students and teachers—that forms a basis for instruction.
5. **Relevant curriculum and instructional practices.** Teachers capitalize on students’ cultural backgrounds to develop challenging curriculum and instructional practices that are relevant to students’ lives.

These five principles were incorporated into new culturally responsive curriculum units (CRCUs) designed by Lab staff and Minority Fellows for this pilot test. All lessons in the fourth- and ninth-grade units include instructional scaffolding that (1) builds on students’ prior knowledge and experience; (2) incorporates their personal and cultural perspectives; and (3) includes instructional materials and approaches that support students’ intellectual, social, emotional, and political growth. The lessons follow a template that reflects Vygotsky’s sociocultural approach to learning—they begin with activities to recruit students’ interest in the concepts to be learned and suggest methods for teachers to communicate high expectations for student learning. Tasks and expectations for performance and assessment are to be explained and clarified for students before they began work. As students work to complete the assignments, the lesson plans provide teachers with activities and suggestions for guiding learning, supporting struggling students, and incorporating enrichment and extended learning opportunities. As teachers assess student work in each lesson, they are to adjust instruction to meet students’ learning needs as they progress through the unit.

The lessons within the CRCUs comply with the West Virginia Content Standards and Objectives (CSO) in the subjects of math, science, social studies, language arts, and art. Furthermore, current teaching approaches and paradigms, such as scientific inquiry, problem solving, values clarification, higher-level thinking skills (Bloom, 1956), and teaching/learning styles (Butler, 1984; Gregorc, 1982) are woven into the lessons. Isolated cooperative learning lessons are not supplied; rather, the design and approach of the curriculum as a whole are intended to be cooperative in nature.

Fourth Grade: Organisms, Organisms Everywhere!

The fourth-grade unit—*Organisms, Organisms Everywhere!*—is designed to help students to grasp how humanity fits within all life forms on Earth. The goals of the unit are to teach students how to recognize and identify ecosystems and environments, within their locale and beyond, in order to understand how every living organism is connected to all others. Students are also to learn how organisms depend on each other for quality of living and survival. At the end of the unit, students should be able to identify how waste, neglect, and development affect present and future cultures, environments, and organisms.

Ninth Grade: Unraveling the Details of the Past

The goal of the ninth-grade unit—*Unraveling the Details of the Past*—is to help educators convey to students that archaeology deepens our understanding of both past and present cultures. This unit is designed to communicate students' relationship to various cultures by critically examining their own environment. Educators are encouraged to use local data and examples, such as mock digs and dioramas where possible. In addition, students are asked to compare their own life experiences with their heritage through real-life examples. As a result of this unit, students should gain a better sense of the interconnected nature of life and cultures. They also should develop an appreciation of archaeology by discovering similarities and differences among cultures through active inquiry, excavation, and practice.

Pilot Test Evaluation Questions

The overarching research questions were as follows: (1) How did teachers who implemented the CRCUs behave? and (2) How could the CRCUs be improved? Specific, related questions include the following:

- Did the classroom instructional behaviors of teachers implementing the culturally responsive units differ from those of teachers in matching schools not implementing the units?
- Did the perceptions and attitudes toward responsive pedagogy, antidiscriminatory practices, high expectations, and differentiated instruction of teachers implementing culturally responsive units differ from those of teachers not implementing the units?
- Was the design of the units workable? Were unit objectives and instructions clear? Were the unit lessons and activities organized clearly and logically? Were the units culturally responsive? Was unit content appropriate?

METHODS

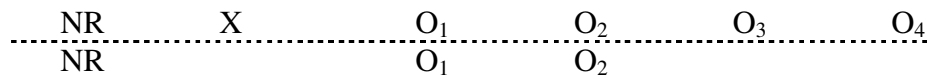
This section describes the perspective, design, time frame, sampling, data collection, and analytic procedures used in this evaluation.

Perspective

This study used both quantitative and qualitative methods. Worthen, Sanders, and Fitzpatrick (1997) note that observations can provide information about context and implementation.

Design

The research design used in this project was pre-experimental in nature. According to Shavelson (1996), an intact-group comparison design is appropriate for collecting pilot test data, as it may “provide useful insights that can be incorporated into other research designs” (p. 25). Gay (1996) concurs that such a comparison is suitable for exploratory studies. The following diagram displays a visual representation of this intact-group comparison design:



Note: NR = Non-random assignment; X = Culturally Responsive Curriculum Units (CRCUs); O₁ = SSOS-R (twice per unit-week); O₂ = MSCI; O₃ = CUER; O₄ = teacher interviews; see Appendix A for further instrument details.

Time Frame

This study was conducted during late spring 2005.

Sampling

The KCS superintendent selected two elementary schools; two fourth-grade teachers from one school would pilot the CRCU and two fourth-grade teachers from the other school were designated as matching teachers. Two high schools were also selected to participate. A few ninth-grade teachers from one school were asked pilot the CRCU and one teacher from the other

school was designated as a matching teacher. Schools that participated in the pilot testing of previous CRCUs were excluded from this pilot test.

Data Collection

Four instruments were employed in this project. Three instruments (the observation system, the teacher survey, and the curriculum unit evaluation rubric) had already been developed, refined, and validated by AEL/Edvantia staff. The fourth instrument (an interview protocol) was developed specifically for this project. KCS assisted with the administration of the instruments and the AEL/Edvantia project director oversaw all data collection. All data collection plans were approved by Edvantia's Institutional Review Board (see Appendix B for informed consent procedures).

Classroom Observations

Teachers from both CRCU and matching groups were observed in the late spring of 2005 using the Special Strategies Observation System—Revised (SSOS-R) classroom observation system. The SSOS-R, which was built on established and systematic classroom observation protocols, has been refined by Lab staff and used successfully in a number of projects. Appendix A further describes the SSOS-R and the various instruments it comprises. Lab staff completed the classroom observations during the time pilot teachers implemented a CRCU. Matching teachers were observed twice a week during the weeks roughly corresponding to the teaching of the CRCU.

Survey

Pilot test teachers were asked to complete the Measure of School Capacity for Improvement (MSCI). The MSCI, a 58-item, seven-scale, self-report instrument, measures the degree to which professional staffs believe their schools possess the potential to become high-performing learning communities. A complete description of the MSCI is contained in Appendix A. The MSCI has been validated (Riffle, Howley, & Ermolov, 2004; Copley, Meehan, Howley, & Hughes, 2005), and the seven subscales have been confirmed in a large-scale, national study (Hughes, Copley, Howley, & Meehan, in press). Internal consistency reliabilities are acceptable, attaining a Cronbach alpha of .97 for the overall instrument and alphas ranging from .77 to .94 for the subscales. The instrument has also been used in other projects (e.g., Hughes et al., 2004).

Unit Feedback

Two primary mechanisms were planned to obtain feedback from teachers who pilot tested the units—an evaluation rubric and a clarifying interview.

Curriculum Unit Evaluation Rubric (CUER). This evaluation rubric, developed and tested with the CRCUs used in the Pilot Schools Project, requires teachers to rate the quality of the instructional units on five dimensions: format, organization, clarity, substance, and cultural responsiveness (see Appendix A).

Teacher interviews. A semistructured interview protocol was to be developed to supplement and clarify MSCI and CUER data. Participating CRCU teachers were to be interviewed after they had completed the MSCI but before the end of the school year.

Data Analysis

Data were analyzed by data collection instrument and by components within those instruments. Because the response rates were low, the underlying assumptions of statistical analyses were not met. Instead, summaries of responses are presented to provide context for the discussion of lessons learned.

FINDINGS

Results will be presented by grade level and by data collection instrument.

Fourth Grade

Fourth-grade teachers participated in most facets of the study—they allowed classroom observations, returned surveys, and provided feedback on the rubric. However, a follow-up interview was not conducted because there was little variance among the responses or no ambiguity requiring clarification.

Classroom Observations

Six hours of classroom observations were collected, including 2 hours of pilot-teacher observations and 4 hours of matching-teacher observations. Data collected via the Special Strategies Observation System-Revised (SSOS-R) were analyzed by instrument section: QAIT¹ Assessment of Classroom, Standards Performance Continuum (SPC), and Classroom Environment and Resources Checklist (CERC).

QAIT. Descriptive statistics of each QAIT question were calculated. The 40 items were listed on a Likert-type scale, which ranged from 1 (*unlike this class*) to 5 (*like this class*).

The pilot teachers scored highest on providing immediate and corrective feedback as well as using extrinsic academic incentives such as making students accountable and guiding partial responses (means of 5.00 with *SDs* of 0.00 each). Pilot teachers scored lowest on items such as using in-class ability grouping and using extrinsic academic and behavioral incentives such as tokens and rewards, small groups with individual incentives, tokens and rewards for improvement, and group contingencies.

The matching teachers scored highest on guiding partial responses (mean of 4.25 and *SD* of 0.58). They scored lowest on using in-class ability grouping, providing individualized instruction, and using extrinsic rewards such as tokens and rewards (means of 1.50 and *SDs* of 1.00 each).

The 40 QAIT items were then grouped into four main categories: quality of instruction, appropriate level of instruction, incentives for learning, and use of time. Pilot teachers' averages on quality of instruction, appropriate level of instruction, incentives for learning, and use of time categories range from 2.97 to 3.96, while matching teachers' averages range from 2.57 to 3.22.

¹ QAIT stands for Quality of Instruction, Appropriate Level of Instruction, Incentives for Learning, and Use of Time.

SPC. The Standards Performance Continuum (SPC) is a classroom observation rubric that yields a quantifiable measure of five pedagogy standards—joint productive activity, language and literacy development, contextualization, challenging activities, and instructional conversation. Five levels of enactment are (a) Not Observed, (b) Emerging, (c) Developing, (d) Enacting, and (e) Integrating. These levels are seen as a continuum of enactment for each of the five standards. Higher numerical ratings correspond to higher levels of enactment.

Pilot and matching teachers did not significantly differ on any of the five pedagogy standards measured on the SPC.

CERC. The Classroom Environment and Resources Checklist (CERC) assesses the presence or absence of indicators of good classroom environments as well as the visibility and usage of a variety of instructional resources.

Pilot teachers tended to incorporate a higher number of indicators of positive classroom environments, including posting classroom rules; displaying student work; having no distracting internal or external noises; and having an open, risk-free environment. Matching teachers had more comfortable ventilation or temperature and greater use of nonsexist materials.

The visible resources also varied between pilot and matching teachers. More instructional resources (i.e., textbooks, worksheets, classroom libraries, reference materials, maps and globes, and audio resources) were both visible and used by pilot teachers than by matching teachers. However, when resources were visible in the classrooms of matching teachers, they tended to use them more often than pilot teachers. Resources observed included textbooks, chalkboard, equipment, television, and video resources.

Survey

In all, three teachers responded to the MSCSI. Two respondents were pilot teachers; the third respondent was a matched teacher. Both pilot teachers were regular classroom teachers. One possessed a bachelor's degree, and the other had earned a master's degree plus 30 or more additional educational credits. They had been teaching for 2 and 18 years, respectively, and the pilot teachers had been teaching at their current schools for 1 and 8 years. Both pilot teachers had been teaching their current subjects for their entire teaching careers, and they had been teaching their current grades for 1 and 5 years. Because only one matched teacher responded to the MSCSI, his/her demographic information will not be presented in order to protect the respondent's confidentiality.

The MSCSI assesses the perceptions of professional staff members regarding how well prepared their school is for undertaking improvement projects successfully and thereby becoming a high-performing learning community. The 58 items of the MSCSI compose seven subscales (defined in Appendix A). Professional staff offer their responses to each item on a six-point, Likert-type scale with 1 representing *not at all true* or *never true* and 6 representing *completely true* or *always true*. Results are presented in the form of item means for each

subscale; higher subscale means indicate greater capacity for improvement (e.g., more readiness to implement improvement programs).

Because so few members of each teacher group responded to the MSCCI, researchers cannot make comparisons between the two groups. However, all three teachers rated their schools' capacity to improve with at least moderate (3.25) to positive (5.17) responses across all seven subscales: Equity in Practice, Expectations for Student Performance, Differentiated Instruction, Improvement in Program Coherence, Peer Reviewed Practice, Coordinated Curriculum, and Technical Resources.

Unit Feedback

Two fourth-grade teachers responded to the Curriculum Unit Evaluation Rubric (CUER). The pilot teachers implemented the units so near the end of the school year that, once the CUER feedback was received, the summer break had already begun. Also, a high percentage of the item responses were rated as *exceptional* and only 11 of the 88 item responses were rated as *acceptable*; no items were rated *improvement recommended*. A follow-up interview was not conducted.

Ninth Grade

Participation in the ninth-grade unit pilot test was quite limited; teachers allowed a total of three 1-hour classroom observations. None of the ninth-grade teachers (either pilot or matching) participated in any further data collection activities.

Classroom Observations

Two ninth-grade pilot test teachers and one ninth-grade matching teacher, all from different schools, allowed classroom observations, which amounted to three 1-hour observations of ninth-grade classrooms throughout this project.

Data from observations using the Special Strategies Observation System-Revised (SSOS-R) were analyzed by instrument section—QAIT, SPC, and CERC. Two pilot-school observations and one matching-school observation were made.

QAIT. Descriptive statistics of each QAIT were calculated. The 40 items were listed on a Likert-type scale, which ranged from 1 (*unlike this class*) to 5 (*like this class*).

The pilot teachers, on average, scored highest on using individualized instruction and using extrinsic academic incentives such as homework checks (means of 5.00 and *SDs* of 0.00

each). They scored lowest, on average, in using extrinsic academic incentives such as implementing small groups with individual incentives and group contingencies.

The matching teacher scored fairly high on many items, including organizing information in an orderly way, noting transitions to new topics, using many vivid images and examples, exhibiting enthusiasm, clearly specifying objectives (on all subitems), and using extrinsic academic incentives such as holding students accountable, homework checks and communication of high expectations (means of 4.00 and no *SDs* because there was only one observation).

The 40 items were then grouped into four main categories: quality of instruction, appropriate level of instruction, incentives for learning, and use of time. The pilot teachers' category means ranged from 3.08 to 4.33; the matching teachers' category means ranged from 2.43 to 3.75.

SPC. The Standards Performance Continuum (SPC) is a classroom observation rubric that yields a quantifiable measure of five pedagogy standards—joint productive activity, language and literacy development, contextualization, challenging activities, and instructional conversation. Five levels of enactment are (a) Not Observed, (b) Emerging, (c) Developing, (d) Enacting, and (e) Integrating. These levels are seen as a continuum of enactment for each of the five standards. Higher numerical ratings correspond to higher levels of enactment.

The pilot teachers scored highest on average for challenging activities (mean of 2.5 and *SD* of 0.71) and lowest on average for language and literacy development. The matching teacher received a score of 1.0 on the instructional conversation scale; the remaining category means were 0.00.

CERC. The Classroom Environment and Resources Checklist (CERC) assesses the presence or absence of indicators of good classroom environments as well as the visibility and usage of a variety of instructional resources.

Pilot teachers had classrooms with higher percentages of observed indicators for good classroom environments, including use of nonsexist materials, posted classroom rules, posted assignments, and on student work displays. Pilot and matching teachers both had classrooms with adequate lighting, comfortable ventilation/temperature, lack of distracting noises (either internal or external in nature), and an open, risk-free environment.

The visible or used instructional resources varied by group. Pilot teachers had classrooms with higher percentages of visibility than matching teachers on worksheets/activity sheets, science and lab tables, student manipulatives, and audio or video resources. Matching teachers had classrooms with higher percentages of visibility than pilot teachers on having a classroom library, instructional aids/props, and overhead projectors. However, the matching teachers did not use any of the resources observed. Half of the pilot teachers actually used the resources that were visible.

Survey

No ninth-grade teachers responded to the survey in either the initial administration or in any follow-up attempts.

Unit Feedback

No ninth-grade teachers responded to the CUER. Therefore, no data existed for analysis; neither could any clarification questions (which would have been based on the rubric responses) be asked in follow-up interviews. Therefore, no interviews were conducted.

DISCUSSION

This section presents a discussion of lessons learned and offers recommendations for further product development.

Lessons Learned

Although sample sizes were small, useful lessons learned and general indicators of unit quality were gleaned. Because pilot test data, by design, are not intended for generalization, the results of this study fit its original purpose—to gather information regarding the quality and workability of units within the intended setting of use.

The lessons learned while conducting this pilot test center on two areas: participation and unit quality and workability.

Participation

One lesson learned was that volunteer pilot teachers' schedules for unit implementation and data collection may not agree with that of the researchers. Pilot teachers may not have had any incentive to implement the unit or to participate in data collection activities. When they did implement a unit, there was poor coordination among the pilot teachers and the researchers. Such scheduling problems led to the inability to follow up with additional data collection activities.

Unit Quality and Workability

Fourth-grade teachers provided very positive feedback regarding the curriculum units. Ninth-grade teachers did not provide any feedback regarding the unit quality. Hence, additional feedback is necessary to gain better insight into the feasibility and workability of the ninth-grade units.

Classroom observation data support the favorable feedback received from the fourth-grade teachers. Pilot teachers' behaviors (such as using in-class ability grouping, holding students accountable, and guiding partial responses) show that they follow the principles of culturally responsive instruction. Pilot teachers tended to provide students with incentives for learning and a positive learning environment. They also possessed and used instructional resources better than their matching counterparts. This supports the notion that implementation of the curriculum units may be associated with more positive behaviors. However, there cannot

be any causation implied; the difference in positive behaviors exhibited may arise simply because the teachers are following a prescribed unit.

The data collectively provide one perspective--teacher behaviors and ratings--of how culturally responsive the units were. Another indicator of responsiveness would be to ask students whether or not they thought that (1) teachers communicated high expectations, (2) the units were responsive to their cultures, (3) they felt in control of academic discourse and (4) the teachers capitalized on their cultural backgrounds. Such information could help illuminate the underlying qualities that would enhance the units and their effectiveness.

Recommendations

The researchers recommend further testing via a field test, in the intended settings of use and in multiple sites, to assess the workability and effectiveness of the units. Also, even if teachers rate the units positively, additional constructive feedback could be useful. No such opportunity was strongly encouraged on the rubric used in this study.

To encourage high levels of participation and response rates, researchers might obtain written agreements with the teachers who will be providing feedback and, perhaps, provide some token remuneration (e.g., honoraria or complimentary classroom resources). These strategies would hold teachers accountable for implementing the units and for adhering to a time frame that fosters data collection within a reasonable field test period.

Random assignment and random selection of participating teachers would also strengthen future research designs, thus removing influences of self-selection or forced implementation.

Finally, it would be interesting to continue this study to see if teachers' perceptions of their schools' capacity to improve increase as they continue to implement culturally responsive curriculum units.

REFERENCES

- Bloom, B. (1956). *Taxonomy of educational objectives, Handbook One: The cognitive domain*. New York: David McKay.
- Butler, K. (1984). *Learning and teaching style in theory and practice*. Columbia, CT: The Learner's Dimension.
- The College Board. (1999). *Projected special context for education of children: 1990-2015*. New York: College Board Publications.
- Copley, L. D., Meehan, M. L., Howley, C. W., & Hughes, G. K. (2005, April). *Second field test of the AEL Measure of School Capacity for Improvement*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.
- Contract Research Corporation. (1975). *Guidelines on pilot and field testing of curriculum products*. Belmont, MA: Author.
- D'Amico, J. (2001, Spring). A closer look at the minority achievement gap. *ERS Spectrum*, 4-10.
- Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of Teacher Education*, 53(2), 106-116.
- Gay, L. R. (1996). *Educational research: Competencies for analysis and application* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Gregorc, A. (1982). *An adult's guide to style*. Maynard, MA: Gabriel Systems, Inc.
- Haycock, K. (2001). Closing the achievement gap. *Educational Leadership*, 58(6), 6-11.
- Hughes, G. K., Copley, L. D., Howley, C. W., & Meehan, M. L. (in press). *Measure of school capacity for improvement: Users' manual and technical report*. Charleston, WV: Edvantia.
- Hughes, G. K., Cowley, K. S., Copley, L. D., Finch, N. L., Meehan, M. L., Burns, R. C., et al. (2004). *Effects of a culturally responsive teaching project on teachers and students in selected Kanawha County, WV, schools*. Charleston, WV: AEL.
- Johnston, R., & Viadero, D. (2000, March 15). Unmet promise: Raising minority achievement. *Education Week*. Retrieved August 30, 2002 from <http://www.edweek.org>
- Ladson-Billings, G. (1994). *The dreamkeepers: Successful teachers of African American children*. San Francisco: Jossey-Bass.

McIntyre, E., Roseberry, A., & Gonzalez, N. (Eds.). (2001). *Classroom diversity: Connecting curriculum to students' lives*. Portsmouth, NH: Heinemann.

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

Riffle, M. J. S., Howley, C., & Ermolov, L. D. (2004, April). *Measure of school capacity for improvement (MSCI): Early field test findings*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Shavelson, R. J. (1996). *Statistical reasoning for the behavioral sciences* (3rd ed.). Boston: Allyn & Bacon.

West Virginia Department of Education. (n.d.) 2001-2002 No Child Left Behind Report Card. Retrieved March 31 2003, from <http://wveis.k12.wv.us/nclb/pub/pickinfo.asp>

Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (1997). *Program evaluation: Alternative approaches and practical guidelines* (2nd ed.). New York: Longman.

APPENDIXES

Appendix A

Data Collection Instrument Descriptions

Special Strategies Observation System-Revised (SSOS-R)

This observation system consists of three distinct data collection protocols, developed and employed in prior research and evaluation studies (Nesselrodt & Schaffer, 1993), refined by Edvantia and Kentucky Department of Education staff, and converted to a scannable format. The three instruments make up the Special Strategies Observation System-Revised (SSOS-R), 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-R is a viable system 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 (Sullivan & Meehan, 1983); the resultant triangulation of data sources further documents the veracity of the data collected. Each instrument that makes up the SSOS-R is described below.

Classroom Observation Form (COF). The instrument is a combination observation form that is best described as a category system, with all low-inference items, and includes multiple coding procedures. It is based on the Classroom Activity Record designed by Evertson and Burry (1989) and the Stallings Observation System (Stallings, 1980). The top page of the form collects typical demographic information, such as the school, observer, date, number of adults and students in class, subject being observed, and type of class. 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 8 minutes of class time. The first minute per page 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 7 minutes per page focus specifically on the target student and include coding 1 of 27 discrete activities for each minute.

QAIT Assessment of Classroom (QAIT). 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, Incentives for Learning, and Use of Time*. This particular instrument was developed by Robert Slavin (1987, 1989) and is based on his research and that of John B. Carroll (1963, 1989). Fitting on two 8½" x 11" 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 is completed at the end of each observation session.

Classroom Environment and Resources Checklist (CERC). This instrument is a low-inference, simple coding sign system. Printed on the front of one 8½" x 11" sheet, it contains 14 classroom attributes that are coded either as *present* or not *present*, such as adequate lighting, use of multiracial materials, and posted assignments. 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 visible, observers indicate whether they are used during the observation. This instrument is to be completed at the end of each observation session.

Standards Performance Continuum (SPC). This instrument is a rubric used to quantify the implementation of the Standards for Effective Pedagogy (Tharp, Estrada, Dalton, & Yamauchi, 2000). It is best described as a high-inference, simple coding, rating device. The SPC contains five standards labeled Joint Productive Activity, Language and Literacy Development, Contextualization, Challenging Activities, and Instructional Conversation. Each standard is rated on a Likert-type response scale of 0 to 4 (*not observed* to *integrating*). This instrument is to be completed at the end of each observation.

Measure of School Capacity for Improvement (MSCI)

The Measure of School Capacity for Improvement (MSCI) is a 58-item questionnaire that is administered to a school's professional staff—administrators, teachers, teachers' aides, librarians, counselors, and any other staff who have significant contact with students and parents. The instrument measures professional staff members' perceptions of how their school is faring in seven different areas related to capacity for improvement. Staff members use a 6-point Likert-type scale to state how true items are of their school or the frequency with which items are true for their school. Completing the MSCI requires approximately 20 minutes.

The 58 items of the MSCI compose seven subscales that encompass important aspects of capacity for improvement. Subscales, which have been validated in a large-scale, nationwide U.S. study, are composed of varying numbers of items. The MSCI and all its subscales possess strong reliability (Cronbach alphas from .77 to .94 for the seven subscales, .97 for the overall instrument) and have demonstrated validity.

Equity in Practice. The 15 items that compose the *Equity in Practice* scale assess equitable practices in the school, specifically responsive pedagogy and anti-discriminatory practices. This scale examines the school's capacity to create an atmosphere of tolerance, cultural awareness, and equity for all learners.

Expectations for Student Performance. The *Expectations for Student Performance* scale contains 11 items. The items assess staff members' expectations of the students and their beliefs that all students can perform well academically.

Differentiated Instruction. *Differentiated Instruction*, formed by 11 items, addresses instructional practices and strategies for reaching students of diverse learning needs. The scale focuses on using or modifying instructional practices to be effective with students of all types.

Improvement Program Coherence. The *Improvement Program Coherence* scale is composed of 9 items pertaining to the extent to which improvement initiatives and efforts at a school are coordinated. The items focus on the coordination of improvement programs or initiatives with existing initiatives and with school improvement goals. Items also focus on school-level support of and for improvement initiatives.

Peer-Reviewed Practice. The 4 items composing the *Peer-Reviewed Practice* scale explore the observation and review by staff of their peers' work. All items assess the extent to which professional staff in a school observe the work of their colleagues and give or receive relevant feedback about their performance.

Coordinated Curriculum. The *Coordinated Curriculum* scale, composed of 4 items, addresses the coordination of curriculum within and across grade levels at the school.

Technical Resources. The 4 items composing the *Technical Resources* scale concern instructional resources and materials, including whether staff possess or have immediate access to adequate materials and resources to achieve instructional objectives.

Curriculum Unit Evaluation Rubric (CUER)

The Curriculum Unit Evaluation Rubric (CUER) was developed for the Pilot Schools project in 2004. Because curriculum unit evaluation tools were not available, Edvantia staff drew from various sources, including checklists, guidelines, and internal expertise to develop the instrument. The rubric assesses curriculum unit format, organization, clarity, substance, and cultural responsiveness. Seventeen items address curriculum format, organization, and clarity; another 15 items focus on the substance of the curriculum units, and the final 11 items assess the cultural responsiveness of the units. Units are rated on each criterion using a 3-option response scale indicating the extent to which the curriculum unit satisfies each item. Response options are as follows: *Exemplary*, worth 3 points and reflecting a commendable unit that is worthy of imitation for that criterion; *Acceptable*, worth 2 points and reflecting a satisfactory unit that is adequate to meet requirements; and *Improvement Recommended*, worth 1 point and reflecting an unacceptable unit that requires revision on that criterion. Points for each of the three major categories (Format/Organization/Clarity, Substance, and Cultural Responsiveness) are then tallied and assigned weights in this manner: 1.0 for Format/Organization/Clarity; 1.5 for Substance; 2.0 for Cultural Responsiveness. A final tally of points across the three categories yields an overall score for each curriculum unit, ranging from 63.5 to 190.5. Curriculum units are then classified according to the total number of points: *Exemplary* (90.5 –148.5), *Acceptable* (48.0 – 106.0), or *Improvement Recommended* (63.5 – 105.5). A unit receiving ratings of *Acceptable* for all items would receive a score of 127.

References for Appendix A

- Carroll, J. (1963). A model of school learning. *Teachers College Record*, 64, 723-733.
- Carroll, J. B. (1989). The Carroll model: A 25-year retrospective and prospective view. *Educational Researcher*, 18(1), 26-31.
- Evertson, C., & Burry, J. A. (1989). Capturing classroom context: The observation system as lens for assessment. *Journal of Personnel Evaluation in Education*, 2, 297-320.
- Nesselrodt, P. S., & Schaffer, E. C. (1993, April). *The ISERP programme: A revised classroom observation instrument*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.
- Slavin, R. E. (1987). A theory of school and classroom organization. *Education Psychologist*, 22(2), 89-108.
- Slavin, R. E. (1989). A theory of school and classroom organization. In R. E. Slavin (Ed.), *School and classroom organization* (pp. 3-18). Hillsdale, NJ: Erlbaum.
- 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). *The development, pilot test, and pilot 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.
- Tharp, R. G., Estrada, P., Dalton, S. S., & Yamauchi, L. A. (2000). *Teaching transformed: Achieving excellence, fairness, inclusion, and harmony*. Boulder, CO: Westview Press.

Appendix B

Informed Consent Procedures

Informed Consent Letter Given to Teachers

My name is Dr. Christy Gilchrist. I am a Research and Evaluation Specialist at AEL, Inc., a nonprofit educational research and evaluation firm that is working with your school district. Joe Holloway, who is my research assistant, and other AEL staff are conducting this **research** to inquire into the quality and workability of certain culturally responsive curriculum units, which were developed by AEL staff.

You were identified as a candidate for this research project because either you are implementing the Culturally Responsive Curriculum Units in your classroom or you are serving as a comparison classroom. I invite you to participate in the following research activities:

Classroom observations: We hope to observe you and your class twice, for one hour each time, during the class periods in which you implement the curriculum materials. The purpose is to assess whether our curriculum materials influence your classroom practices—it is not to evaluate you or your students in any way. For instance, we'll document the types and lengths of typical classroom interactions. When an AEL researcher first enters your classroom, he or she will confirm your agreement with these terms prior to collecting data. Once your consent is obtained, the researcher must follow a very strict data collection schedule. Therefore, once he or she begins filling out the bubble sheets, please refrain from most any interaction—it may skew the results.

Surveys: We would also like you to complete two surveys. First, the AEL Measure of School Capacity for Improvement (AEL MSCI) will help us understand your perceptions regarding school climate. The second instrument, called the AEL Curriculum Unit Evaluation (AEL CUER), was developed to systematically solicit your feedback regarding the quality and the workability of our products during a pilot test. We will not collect your name on any of these instruments.

Interviews: We would like to follow up with you after we collect and analyze your feedback regarding the quality and workability of the curriculum units. This will allow us to further explore your perceptions and reactions. Additional details will be provided at that time.

Your requested assistance depends on whether you implement a CRCU. See Table 1 for details.

Table 1: Participation Guidelines

Who	What	When
All Teachers	Classroom Observations	Twice during unit, or during the time CRCU teachers teach a unit
All Teachers	Complete AEL MSCI survey	Any time during the school year
CRCU Teachers	Complete AEL CUER	After teaching the CRCU unit
CRCU Teachers	Interviews	After teaching the CRCU unit and completing the AEL CUER

None of these procedures are deemed to be experimental; rather, the observation and survey instruments are fully developed and tested. Furthermore, the AEL Institutional Review Board has approved these activities and found them to be ethical and in compliance with federal regulation.

Every effort will be made to ensure confidentiality of the data collected, such as using pseudonyms in reporting and keeping the master list of participant names separate from the data. Furthermore, only AEL staff authorized by me will access the data strictly for the purposes of organizing and analyzing the results. Data will be stored on a password-protected computer file for three years past the termination of this project. Your names will not be stored with any of the data collection instruments. Finally, the results from our research activities *will not be shared directly with anyone*, including your peers, principal, or superintendent. Rather, results will be aggregated by grade level and reported without names or other identifying information. Disclaimer: The AEL IRB has the authority to inspect consent records and data files only to assure compliance with approved procedures and with federal regulation.

There are no known risks associated with participating in these research activities that are greater than those ordinarily encountered in daily life. Regardless, you may initially feel a small amount of anxiety as you would with anyone observing you. Should you feel too anxious, please notify a researcher immediately. Information obtained during this study will be held in strictest confidence, with the exception that if a researcher obtains clear evidence of unlawful behavior that could result in physical or mental damage to a minor, the researcher is required by statute to report such evidence to the authorities.

There are no expected *immediate* direct benefits for participating teachers. Nor will there be any compensation given by AEL to participate in these research activities. However, the results will be used to improve AEL Culturally Responsive Curriculum Units, which may enable AEL staff to create more culturally responsive curriculum materials that can be used to close the achievement gap.

If you have any question or concern, please feel free to contact me, Dr. Christy Gilchrist, at 1-800-624-9120 extension 5474 or by e-mail at gilchristc@ael.org. For information on your rights, contact Dr. Merrill Meehan, AEL IRB Chair, 1031 Quarrier Street, 6th Floor, or at 1-800-624-9120 ext. 5432.

Please remember that your participation is completely voluntary. This means that you can discontinue your participation at any time without reprisal or penalty.

