Different Ways of Knowing for the Middle Grades: Cohort 1 Evaluation Report

Steven M. Ross
Barbara Wilson-Relyea
Aaron McDonald
Elizabeth Goldfeder
John Nunnery
Center for Research in Educational Policy

Marco Muñoz
Jefferson County Public Schools

Center for Research in Educational Policy
University of Memphis
325 Browning Hall
Memphis, TN 38152
Toll Free: 1-866-670-6147

June 2004

EXECUTIVE SUMMARY

This report presents results from a planned comprehensive evaluation of the implementation of Different Ways of Knowing (DWoK) for the Middle Grades in "Cohort I" schools. Cohort I consists of 15 schools in California, Kentucky, New York, and Texas that began implementation of DWoK for the Middle Grades in 2001-02, and thus have completed their second year at this time.

The planned comprehensive evaluation is comprised of four sub-studies. The first study is an analysis of results of a teacher survey, developed and administered by WestEd in the four Cohort I states. The second study synthesizes interview results obtained from DWoK for Middle Grades instructional and artist coaches and the team leader in one of the four states (KY). For this second study, the main generic findings across the five schools are summarized. In order to protect the promised confidentiality of the individual schools, the initial report – which contained very specific information about each school - was submitted to Galef as a separate formative evaluation report for internal use. The third study presents results from analyses of student achievement, attendance, and school climate comparing three urban DWoK schools to three matched control schools in the Jefferson County Public School district (KY). The fourth study was an analysis of achievement at a DWoK rural school compared to matched rural control school.

Evaluation Questions:

1. What is the quality of the Different Ways of Knowing model?
2. How well is the model being implemented at the school sites? High/low fidelity?
3. In what ways has the model resulted in improvements in student achievement, teacher practice, and other components of comprehensive school reform?

Study I: Teacher Survey

Method

Teachers were asked to provide impressions of their schools' implementation of DWoK. The survey instrument was organized into six main categories which centered around different components of DWoK — Standards, Instruction, Literacy, Assessment, Leadership, and Reform Implementation. All scales were found to have an $\alpha > .60$ indicating a moderate to high degree of reliability in the response sets.

Results

Overall Sample

A total of 302 teachers from the four states returned completed surveys. The majority of the respondents were from Kentucky (41.4%) and New York (35.4%). Strongest among the positive responses was the extent to which teachers used standards, as well as the confidence the teachers placed in their own ability to provide organized, meaningful, and student specific instruction. Weakest was their confidence in the school administration and DWoK staff to provide them with the time, resources, and expertise to improve their own performance as well as the schools' performance. See Appendix A for summary table of teacher survey results.

Standards and Instruction were two relatively strong implementation areas, revealing teachers' engagement in self-evaluation activities. Literacy responses indicated a moderate degree of strategy use. Assessment responses, like most categories, tended to fall in the middle of the scales. The least positive responses related to allowing students to learn self-assessment skills.

Leadership was an area of concern. Leadership responses were less positive overall than for other categories. Although moderately positive about the leadership

staying abreast of and sharing the goals of the reform plan, the teachers had concerns about of the viability of school and reform administrators intervening in classroom dynamics.

Implementation responses were mixed. Implementation strengths included the commitment of administration to the DWoK goals, implementation of curriculum standards, and the overall fit of the program with the school. Implementation concerns were expressed in the areas of meaningful and effective teacher involvement in the reform process as well as time constraints.

Differences Between States

No large differences in responses are shown between any of the states. Results for Kentucky tended to be more positive than those for other states, especially on Leadership and Implementation. California scored somewhat lower in the Standards, Instruction, and Assessment categories, whereas Texas scored lower in the Literacy and Leadership categories.

Study II: Coaches' Interviews at the Kentucky Middle Schools

Method

Formal structured phone interviews were conducted with the instructional coach, artist-educator coach, and team leader at each of the five Kentucky middle schools implementing DWoK. Of the five Kentucky schools, two were small rural schools, and three were larger urban schools. Areas assessed included DWoK's design components, effectiveness, implementation quality and progress, and perceived impact.

Results

Implementation Progress

- At three schools, implementation has been going well. At two schools, however, DWoK implementation has been either "surface-level" or non-existent.

Teacher Support

- At three of the schools, there appeared to be solid support for DWoK. At two schools, however, teacher support was guarded at best.

Impact on Students

- As specifically noted in four of the five schools, DWoK was reported to have positively impacted student sharing and learning. Additionally, students at two schools were reportedly more enthusiastic about learning as a result of DWoK.

Impact on Teachers

- As noted in four of the five schools, there have been improvements in teacher sharing and/or engagement as a result of DWoK. There also appears to have been some growth in teacher planning and instruction as a result of DWoK. Further, as specifically noted in two of the five schools, teacher relationships appear to have improved.

Study III: Student Achievement Analyses in Urban Schools

Method

Using a rigorous quasi-experimental design, three urban KY DWoK schools were matched to three similar control schools. Control schools were selected based on similarity to DWoK schools according to several school-level and student-level variables. School-level variables included: % of students eligible for free or reduced-price lunch, student mobility, attendance, special education (Exceptional Child

Education, ECE), single-parent households, and the state's Accountability Index (AI). Student-level variables included previous test scores in reading, age, gender, race, free or reduced-price lunch, and single-parent households in the baseline year (2000-2001). For Year 1, seventh-grade KCCT reading student-level data were analyzed via *ANCOVA*, with the 2000-2001 (sixth-grade) CTBS reading "pretest" scores as a covariate. For Year 2, data were analyzed using (a) *MANCOVA*, examining the sixth-grade language arts, reading, and mathematics subtests of the CTBS; and ANCOVA, examining seventh-grade KCCT Reading and Arts & Humanities scores with sixth-grade Stanford Diagnostic Reading Test (SDRT) scores as a covariate.

Results

Year 1: Overall Program Effects and School Comparisons

In Year 1 (2001-2002), three DWoK schools demonstrated a significant overall advantage (*ES* = +0.19) over their matched control schools on seventh-grade KCCT reading. Individual school comparisons lacked comparable power and therefore, although directionally favoring the DWoK schools in each case, were significant only for Pair C. However, the associated effect sizes of +0.18, +0.31, and +0.20, were at least suggestive regarding program impacts. For all three pairs, the CTBS Reading pretest covariate was highly significant.

Year 2: Multivariate Analysis of Covariance on CTBS Subtests

In Year 2 (2002-2003), the DWoK mean was significantly higher than the control group mean on all three CTBS subtests of reading, language arts, and mathematics. The effect sizes associated with the unadjusted means were relatively small in magnitude ranging from +0.10 to +0.13. Those associated with adjusted means were slightly higher, ranging from +0.16 to +0.18.

Year 2: Analysis of Covariance on KCCT Subtests

ANCOVA results on 7th grade KCCT Reading, using fall 2001 Stanford Reading Diagnostic Test (SDRT) scores as the covariate, yielded a significant Program effect [$F(1,1589) = 5.81$, $p = .016$, $ES = +0.17$] favoring DWoK ($M_{adj} = 511.78$, $SD = 35.82$) over control ($M_{adj} = 508.84$, $SD = 32.12$). A parallel ANCOVA conducted on Arts & Humanities scores did not approach significance [$F(1, 1569) = .002$, $p = .962$]. The adjusted DWoK and Control means ($M_{adj} = 503.34$ and 503.50, respectively) were nearly identical.

Attendance Rates

Attendance rates at DWoK schools decreased by about 0.60 of a percentage point from the baseline year to Year 2 of implementation. During this same period; attendance at the control schools decreased by 1.37 points, over twice the rate of decline for DWoK. The 2(program) x 3(year) ANOVA showed the interaction pattern to be significant: although there were no program group differences in 2000-01 (baseline) or 2001-02 (Year 1), the DWoK schools significantly surpassed ($p < .001$, $ES = +0.19$) the control schools in 2002-03 (Year 2).

School Climate and Perceptual Data

Taken as a whole, the data do not support the conclusion that conditions at DWoK schools systematically differed from those at Control schools. Results for DWoK-A and DWoK-B were slightly favorable, while a moderate negative pattern was indicated for DWoK-C.

Study IV: Student Achievement Analyses in a Rural School

Method

The sample included 7th and 8th grade students from 2001 to 2003 in two schools: one DWoK schools, and one comparison school. The comparison school was selected on the basis of similarity to the DWoK schools in terms of the percentage of students eligible for free or reduced-price lunch, student mobility, attendance, special education (Exceptional Child Education, ECE), percentage of single-parent households, and the state's Accountability Index. Three cohorts were present in the sample: 2002 eighth graders, 2003 eighth graders, and 2003 seventh graders. For seventh grade, the KCCT subtests in Reading and Science were administered. For eighth grade, the KCCT subtests in Mathematics, Social Studies, Arts and Humanities, and Practical Living/Vocational Studies were administered.

Results and Discussion

Inferential tests of program effects showed a mixed pattern of results. For the 2003 7th grade cohort, the comparison school had a statistically significantly higher mean in Science, although the effect was small (ES = -0.15), while the DWOK school had a significantly higher mean in Reading (ES = +0.24). Results for the 2002 8th grade cohort favored the DWOK school in Social Studies (ES = +0.18) and Practical Living (ES = +0.43). For the 2003 8th grade cohort, results clearly favored the comparison school. The comparison school had significantly higher mean scores in Mathematics (ES = -0.36), Social Studies (ES = -0.31), and Arts and Humanities (ES = -0.27).

Discussion

This section will summarize the study's findings as they relate to each of the three research questions.

1. What is the quality of the Different Ways of Knowing model?

According to the teacher survey, Standards and Instruction appeared to be strengths. Teachers were highly positive about Standards, relative to the other categories. Leadership appeared to be a concern. Teachers most frequently disagreed about the confidence in their school administration and DWoK staff to provide them with the time, resources, and expertise to improve their own performance as well as the schools' performance.

2. How well is the model being implemented at the school sites? High/low fidelity?

Implementation strengths in all four states centered around the commitment and involvement of the school administration to DWoK goals, alignment with curriculum standards, and the overall fit of DWoK with their school. Implementation weaknesses concerned meaningful and effective teacher involvement in the reform process, as well as time constraints in order to effectively support curriculum integration and de-tracking.

Teachers appeared to be more positive about DWoK implementation at the Kentucky schools than the other states. Still, implementation at the Kentucky schools appears to have differed widely depending on the school. At one of the schools, implementation has been "wonderful." At another school, implementation has also been going well, although teachers are overloaded with the large number of other programs and initiatives. At a third school, implementation has been going very well, despite being difficult at first. At the final two schools, implementation has been either nonexistent or only "surface-level."

3. *In what ways has the model resulted in improvements in student achievement, teacher practice, and other components of comprehensive school reform?*

DWoK appears to have had a positive impact on student achievement in the three urban Kentucky schools included in the achievement analyses, as compared to matched control schools. After the first year of DWoK implementation, the three schools demonstrated a significant overall advantage over their matched control schools on seventh-grade KCCT Reading. Furthermore, on all three sixth-grade CTBS subtests (i.e., math, language arts, and reading) and on seventh-grade KCCT Reading in Year 2, DWoK schools surpassed the matched control schools. Although the effect sizes were modest in size, they still approached levels associated with more established CSR models (Borman et al., 2003). Of further note, attendance rates over the three-year period from 2000-01 to 2002-03 significantly favored DWoK over control schools.

For the rural school, however, inferential tests of program effects showed a mixed pattern of results. Importantly, DWoK was favored in Reading in the seventh grade in 2003 (ES = +0.24), but had a small deficit in Science (ES = -0.15). Results for the 2002 8th grade cohort on the KCCT favored DWoK on several subtests, but favored the comparison school in 2003 on multiple tests. Thus, the rural school achievement results are equivocal at this time.

DIFFERENT WAYS OF KNOWING FOR THE MIDDLE GRADES: COHORT I EVALUATION REPORT

Introduction

This report presents results from a planned comprehensive evaluation of the implementation of Different Ways of Knowing (DWoK[1]) for the Middle Grades in "Cohort I" schools. The evaluation supports the development work by the Galef Institute as part of the "capacity building grant" received from the U. S. Department of Education, ending in 2003. Cohort I consists of 15 schools that began implementation of DWoK for the Middle Grades in 2001-02, and thus have completed their second year at this time. Cohort I sites were established in the states of California, Kentucky, New York, and Texas.

As strongly emphasized in the *No Child Left Behind* (NCLB) legislation, it is critically important to identify effective reform models and interventions on the basis of "what works" in bringing all students to proficiency levels in core subjects. Thus, the planned comprehensive evaluation is comprised of four sub-studies. The first study is an analysis of results of a teacher survey, administered by WestEd in the four Cohort I states. The second study synthesizes interview results obtained from DWoK for Middle Grades instructional and artist coaches and the team leader in one of the four states (KY). For this second study, the main generic findings across the five schools will be summarized. (The original report, containing very specific information about the schools provided by the coaches, is being submitted to Galef as a separate formative evaluation report for internal use, in order to protect the promised confidentiality of the individual schools.) The third study (and also the most critical study) presents results

[1] For brevity, *Different Ways of Knowing* will be abbreviated *DWoK* throughout this report.

from analyses of student achievement comparing three urban DWoK schools to three matched control schools in the Jefferson County Public School district, as well as their implications for model scale-up and further capacity building. The fourth study was an analysis of achievement at a DWoK rural school compared to matched rural control school.

The design and methodology of this report was oriented around the following research questions:

1. What is the quality of the Different Ways of Knowing model?
2. How well is the model being implemented at the school sites? High/low fidelity?
3. In what ways has the model resulted in improvements in student achievement, teacher practice, and other components of comprehensive school reform?

The Different Ways of Knowing Model

Different Ways of Knowing (DWoK) is a multiyear, comprehensive design for teacher quality and school improvement. The original, elementary-grades model was developed in 1989 by the Galef Institute, a nonprofit educational organization dedicated to school reform. As described by Herman (1999), DWoK builds on the "multiple intelligences' of students, to develop their skills in different areas (such as artistic, mathematical, social, language). The model strongly emphasizes thematic units that integrate the learning of basic literacy and mathematics skills with artistic experiences. Specifically, according to Herman (1999), "The curriculum, which is organized around history and social studies, seeks to integrate the arts, literature, science, and math, and technology" (p. 59). DWoK met the rigorous standards to qualify as a New American Schools design and is a member of the Public Education Network.

DWoK for the Middle Grades was developed starting in 2000 under a $13 million contract with the U.S. Department of Education. The model was designed to provide tools, products, and services that produce results for young adolescents. According to the developer, the core assumption is that all students, teachers, and school leaders have the capacity to develop expertise in any given subject or skill, DWoK for the Middle Grades provides learners varied instructional pathways. By integrating the visual, performing, literary, and media arts, the goal is to promote in-depth, creative thinking and content acquisition across all disciplines and motivate students to think critically and gain deeper, long-lasting understandings.

The developer advocates that their products, tools, and services focus on mathematics, reading, and writing literacy, to:

- Support educators to provide experiences for young adolescents that are responsive to their academic, developmental, and social needs
- Support schools and districts to meet the requirements of No Child Left Behind
- Provide a menu of services that reduce achievement gaps for all student groups
- Support state-developed essential components and guidelines for effective middle-grades programs
- Operationalize the recommendations of *Turning Points 2000*
- Influence national conversations around the design of programs that meet the unique needs of young adolescents and the adults who work with them

Specifically with regard to NCLB and AYP goals, DWoK for the Middle Grades focus on teacher quality, leadership, and parent and community engagement in:

- Standards-based planning in curriculum, assessment, and instruction
- Student inquiry and self-directed learning

- Comprehensive schoolwide literacy program, including expert strategies in reading and writing
- Integrating the arts for deeper content learning
- Shared leadership for results
- Organizational structures that support teaching, learning, and a positive school climate

Study I. Teacher Survey Results

Introduction

This report summarizes the results of a survey administered by WestEd to assess teacher reactions to their activities and experiences in implementing Different Ways of Knowing (DWoK) for the Middle Grades. Specifically, teachers in four states — California, Kentucky, New York, and Texas — were asked to provide impressions of their schools' implementation of DWoK. To protect confidentiality of responses for teachers and schools, WestEd collected the surveys from the schools and bundled them by state. Thus, the present analysis is restricted to providing aggregate results for individual states and for the total sample of 302 teachers.

Method

Instrumentation

The survey instrument is divided into six main categories and 17 subcategories. The main categories deal with different components of DWoK — Standards, Instruction, Literacy, Assessment, Leadership, and Reform Implementation. The subcategories are formed by different scales used to record responses (e.g., Likert-type ratings, checklists, frequency of use, etc.). Specifically, within each category there is at least one set of questions using a four-point Likert scale ("Strongly Disagree" to "Strongly Agree") or a

five-point Likert scale ("Never" to "Always" or "Almost Daily"). The second category (Instruction) also includes items using a six-level scale (None to 76-100%). Multiple response dichotomies are included in category three (Literacy). Each choice given is treated as a separate question and the results indicate the proportion of respondents who selected that choice.

Due to the variety of response types, items were grouped into subcategory scales by response type under each of the six main categories. To determine internal consistency, we computed Cronbach's Alpha for each subcategory scale. All scales were found to have an $\alpha > .60$ indicating a moderate to high degree of reliability in the response sets. Note that the subcategories (in italics) in the summary below are labeled by type of scale (e.g., "frequency") and ordinal number within that type. For example "Frequency 1" refers to the *first* set of items in the particular category using a frequency scale, whereas "Frequency 2" refers to the *second* set of such items.

Standards: The extent to which the teacher makes use of specific standards in instruction.

Agreement 1 (14 items, $\alpha = .85$)
Frequency 1 (3 items, $\alpha = .63$)

Instruction: The extent to which actual instruction is carried out with specific goals in mind that connect to DWoK.

Percent of Time 1 (4 items, $\alpha = .65$)
Agreement 1 (6 items, $\alpha = .81$)
Frequency 1 (6 items, $\alpha = .82$)

Literacy: The extent to which the teacher incorporates broad areas of knowledge-based literacy in lessons.

Multiple Response 1 (checklist – 9 choices)
Multiple Response 2 (checklist – 10 choices)
Frequency 1 (Writing Tasks, 3 items, $\alpha = .89$)
Frequency 2 (Reading & Thinking tasks, 6 items, $\alpha = .77$)
Frequency 3 (Comprehension Strategies, 9 items, $\alpha = .94$)

Frequency 4 (Modeling Strategies, 5 items, α= .95)
Multiple Response 3 (checklist – 7 choices)
Agreement 1 (1 item)

Assessment: The extent to which the teacher involved students in self-assessment.

Frequency 1 (Student Self-assessment, 4 items, α= .80)
Frequency 2 (Modeling Assessment, 4 items, α= .83)

Leadership: The extent to which school administration provides leadership aligned with DWoK.

Agreement 1 (9 items, α= .88)
Frequency 1 (4 items, α= .76)

Implementation: The extent to which the Community, Administration and Teaching staff take ownership of DWoK as a process.

Agreement 1 (Administration, 6 items, α= .90)
Agreement 2 (School & Community, 19 items, α= .92)

Results

Respondents

A total of 302 teachers from the four states returned completed surveys. As shown in Table 1, two states, Kentucky and New York, made up the majority of respondents.

Table 1

Different Ways of Knowing for the Middle Grades: Teacher Implementation Survey Returns for the Four States

State	Number	Percent
California	20	6.6%
Kentucky	125	41.4%
New York	107	35.4%
Texas	50	16.6%
Total	302	100.0%

Item Scales

Positive responses differ by type of scale. For the 4-point Likert scale used in Agreement subcategories, a positive response is "Agree" and "Strongly Agree" (values "3" and "4") combined. For the 5-point Likert scales used to measure Frequency subcategories, a positive response is either "Sometimes" combined with "Often" and "Always" or "2-3 times per month" combined with "1-2 times per week" and "Almost daily" (values, "3", "4" and "5"). Positive values on the percentage scale were above 50% (values "5" and "6"). For the following discussion, "Agreement" refers to a positive response and "Disagreement" to a negative response; "Frequently" to a positive response and "Infrequently" to a negative response; "Greater than half the time" to a positive response and "Less than half the time" to a negative response.

As will be described below, overall, positive responses were indicated in a number of different areas. Strongest among these was the confidence the teachers placed in their own ability to provide organized, meaningful, and student specific instruction. Weakest was their confidence in the school administration and DWoK staff to provide them with the time, resources, and expertise to improve their own performance as well as the schools' performance.

Aggregate Results Across States

Category I: Standards

Tables 2 and 3 review item results in the Standards category. This category has the highest overall positive response of the seven categories. Half of the items in the *Agreement 1* scale have a greater than 90% positive rating and two thirds of the items in the *Frequency 1* scale are greater than 90% positive.

Agreement 1 (4-pt. Scale). As shown in Table 2, the *Agreement 1* set has an overall mean of 3.14, indicating generally positive views. As indicated by the percentages, the highest level of agreement occurred on Q2, "All my lessons have specific learning goals" (98.7%, *M*=3.59 out of 4) and Q13, "I analyze student work to determine how I need to modify my instruction" (98.3%, *M*=3.39). A high level of agreement is also seen on items Q8, "I give my students continuous feedback while they are working on projects" (94.6%, *M*=3.32) and Q3, "All the learning goals of my lessons are based on standards" (94.3%, *M*=3.43).

The highest level of disagreement occurred for item Q7, "I have time available to plan standards-based instruction with my colleagues" (57.7%, *M*=2.26). Other items with relatively high disagreement are Q9, "I provide students with the opportunity to critique each other's work" (21.8%, *M*=2.93) and Q4, "I always introduce the standards I will use in my lessons to students" (20.8%, *M*=3.01).

Table 2

Category I: Standards - Agreement 1 (4-pt. Scale): Descriptive Statistics and Item Response Percentages

	Scale Prompt	***N***	***Min***	***Max***	***Mean***	***SD***	***Alpha***
	To what extent do you agree or disagree with the following?	264	2.26	3.59	3.14	0.31	0.85
	Item	**Strongly Agree**	**Agree**	**Disagree**	**Strongly Disagree**		
Q1	I refer to standards when giving feedback to students on the quality of their work.	27.2	61.4	10.4	1.0		
Q2	All my lessons have specific learning goals.	59.7	38.9	1.3	0		
Q3	All the learning goals of my lessons are based on standards.	48.0	46.3	5.0	0		
Q4	I always introduce the standards I will use in my lessons to students.	24.2	54.0	19.1	1.7		
Q5	I use multiple standards within a content area to teach a "big Idea" (core concept).	31.5	59.7	7.4	0.7		
Q6	I use multiple standards across content areas to teach a "big Idea" (core concept).	23.5	60.1	14.1	0.7		

Table 2 (continued)

Category I: Standards - Agreement 1 (4-pt. Scale): Descriptive Statistics and Item Response Percentages

Q7	I have time available to plan standards-based instruction with my colleagues.	5.0	34.9	40.6	17.1
Q8	I give my students continuous feedback while they are working on projects.	36.9	57.7	4.4	0.3
Q9	I provide students with the opportunity to critique each other's work.	17.1	60.7	20.5	1.3
Q10	Teachers at this school spend time developing assessments that are clearly aligned to the state standards.	23.2	57.7	13.1	2.7
Q11	I meet periodically with students to review their work.	21.1	63.1	14.1	1.3
Q12	I use a variety of strategies to diagnose the areas in which my students are not meeting standards.	27.9	65.1	5.7	0.7
Q13	I analyze student work to determine how I need to modify my instruction.	40.3	58.1	1.0	0
Q14	I assess student work based on criteria set by state standards.	28.2	62.1	8.1	0

Frequency 1 (5-pt. Scale). The *Frequency 1* set has an overall mean of 3.82, which is approximately an answer of "Often." The frequencies in Table 3 indicate the highest positive rating for Q16, "Align classroom assessments to standards" (96.0%, *M*=4.11). The most infrequently used strategy was on Q17, "Use backward planning to develop lessons" (15.8%, *M*=3.44).

Table 3

Category I: Standards - Frequency 1 (5-pt. Scale)

	Scale Prompt	***N***	***Min***	***Max***	***Mean***	***SD***	***Alpha***
	This school year, how often did you do the following?	294	3.44	4.11	3.82	0.34	0.65
	Item	**Never**	**Not Very Often**	**Some-times**	**Often**	**Always**	
Q15	Present standards to students in language they can understand.	1.7	4.4	24.5	40.6	28.5	
Q16	Align classroom assessments to standards?	0.3	2.3	18.1	43.3	34.6	
Q17	Use backward planning to develop lessons.	3.4	12.4	34.6	34.2	13.8	

Category II: Instruction

Tables 4 to 6 review results for the Instruction category. Responses in this category were positive, but somewhat less so than in the Standards category just described.

Percent of Time (6-pt. Scale). As indicated in Table 4, the Percent of Time category has an overall mean of 5.12 indicating that on average, greater than 50% of lessons addressed clearly defined and standards-based topics. Consistent with Category I, the highest proportion of respondents (Q18, 88%, *M*=5.50) indicated that 50% or more of their lessons "were clearly aligned with state standards." The greatest variability and the largest proportion of negative responses occurred for Q21, "Focused on a Big Idea" (37.9% < 50, *M*=4.65).

Table 4

Category II: Instruction – Percent of Time (6-pt. Scale)

	Scale Prompt	*N*	*Min*	*Max*	*Mean*	*SD*	*Alpha*
	Consider the lessons you have taught this year. What percentage of your lessons:	**286**	**4.65**	**5.5**	**5.12**	**0.40**	**0.65**
	Item	**None**	**1-10**	**11-25**	**26-50**	**51-75**	**76-100**
Q18	Were clearly aligned with state standards?	0	0.3	1.7	8.1	26.8	61.4
Q19	Focused on basic facts, concepts and procedures related to a topic?	0	0.3	3.0	12.1	25.5	56.7
Q20	Focused on studying a topic in depth?	0.3	0.3	8.1	20.1	34.6	32.9
Q21	Focused on a "big idea"?	2.7	6.0	8.1	21.1	27.9	31.9

Agreement 1 (4-pt. Scale). As shown in Table 5, the *Agreement 1* set has an overall mean of 3.14, indicating generally positive views. The highest level of agreement occurred on Q25, "I plan lessons that provide opportunities for continuous learning" (96%, *M*=3.31). There was also general agreement for Q23, "I plan lessons that build on students' knowledge and skill through research and collaboration" (89.6%, *M*=3.16) and

Q24, “I plan lessons that develop students' expertise, deep understanding, and presentation skills” (89.3%, *M*=3.19). There was some disagreement in response to Q22, “I plan lessons based on what students already know” (23.8%, *M*=2.87).

Table 5

Category II: Instruction – Agreement 1 (4-pt. Scale)

	Scale Prompt	*N*	*Min*	*Max*	*Mean*	*SD*	*Alpha*
	To what extent do you agree or disagree with the following?	280	2.87	3.31	3.14	0.14	0.81
	Item	**Strongly Agree**	**Agree**	**Disagree**	**Strongly Disagree**		
Q22	I plan lessons based on what students already know.	13.8	58.7	21.5	2.3		
Q23	I plan lessons that build on students' knowledge and skill through research and collaboration.	24.2	65.4	5.7	1.7		
Q24	I plan lessons that develop students' expertise, deep understanding, and presentation skills.	27.5	61.7	7.4	0.7		
Q25	I plan lessons that provide opportunities for continuous learning.	33.2	62.8	1.3	0.7		
Q26	In my classroom, the arts are used to deepen understanding of new learning.	27.9	56.7	11.4	1.3		
Q27	In my classroom, the arts are used to help students demonstrate their learning.	29.9	56.7	8.4	1.0		

Frequency 1 (5-pt. Scale). As shown in Table 6, the *Frequency 1* set has an overall mean of 3.50, indicating that the activities were accomplished “Often.” The most frequently used activities in this set are Q31, “Explain to students how a current lesson is linked to prior lessons and future lessons” (94.3%, *M*=4.07) and Q33, “Provide students with examples of exemplary work to clarify expectations” (91.3%, *M*=3.80). More infrequently occurring items were Q28, “Use the Backwards Planning Template when planning instruction” (31.2%, *M*=2.94) and Q30, “Use guiding questions with the Questioning Triangle to relate the "big idea" to lesson content” (28.5%, *M*=3.07).

Table 6

Category II: Instruction – Frequency 1 (5-pt. Scale)

	Scale Prompt	*N*	*Min*	*Max*	*Mean*	*SD*	*Alpha*
	This school year, how often did you do the following?	282	2.94	4.07	3.50	0.48	0.82
	Item	**Never**	**Not Very Often**	**Some-times**	**Often**	**Always**	
Q28	Use the Backwards Planning Template when planning instruction.	7.4	23.8	40.6	18.8	7.0	
Q29	Use generalizations about "big ideas" to relate them to lesson content.	3.7	14.8	35.2	36.2	7.4	
Q30	Use guiding questions with the Questioning Triangle to relate the "big idea" to lesson content.	7.4	21.1	34.2	26.2	8.1	
Q31	Explain to students how a current lesson is linked to prior lessons and future lessons.	0.7	3.0	15.8	48.3	30.2	
Q32	Use guiding/building questions to scaffold student learning.	1.3	7.7	21.5	43.0	23.5	
Q33	Provide students with examples of exemplary work to clarify expectations.	1.0	4.7	26.5	44.0	20.8	

Category III: Literacy

Tables 7 -14 review items in the Literacy category. Responses in this category indicated a moderate degree of strategy use, reflecting the great investment in time that such instruction requires. This finding is consistent with the general disagreement to Category I (Standards) questions relating to time available and to Category II (Instruction) questions relating to student-individualized instruction.

Multiple Response 1 (checklist – 9 choices). Table 7 shows the proportion of respondents who checked each choice in response to Q34, “What types of materials do you assign your students to read?” The most frequently selected choices were Q34a, “Chapters in textbooks” (71.5%) and Q34f, “Non-fiction books or articles” (70.4%). The least frequently selected choices were Q34h, “Vocational, work-related, or other functional print materials” (39.1%) and Q34d, “Novels” (41.9%).

Table 7

Category III: Literacy – Multiple Response 1 (Checklist)

	Prompt	
	What types of materials do you assign your students to read?	
	Item	**%Checked**
Q34a	Chapters in textbooks	71.5
Q34b	Short Stories	57.0
Q34c	Poetry	50.0
Q34d	Novels	41.9
Q34e	Encyclopedias and other reference materials	59.9
Q34f	Non-fiction book or articles	70.4
Q34g	Plays, film, or television scripts	45.8
Q34h	Vocational, work-related, or other functional print materials	39.1
Q34i	Web-based resources (internet)	57.0

Multiple Response 2 (checklist – 10 choices). Table 8 shows the proportion of respondents who checked each choice in response to Q35, "In general, what types of writing do you assign your students?" The most frequently selected choices were Q35j, "Note taking" (74.5%) and Q35a, "Worksheet exercises (e.g., phonics, vocabulary, grammar, etc.)" (71.0%). The least frequently selected choice was Q35i, "Plays; film; or television scripts" (27.3%).

Table 8

Category III: Literacy – Multiple Response 2 (Checklist)

	Prompt	
	In general, what types of writing do you assign your students?	
	Item	**%Checked**
Q35a	Worksheet exercises (e.g., phonics, vocabulary, grammar, etc.)	71.0
Q35b	Practice sentences and paragraphs	47.2
Q35c	Newspaper and magazine articles	57.3
Q35d	Themes and essays	51.4
Q35e	Research papers	50.3
Q35f	Brief expository or informational pieces (e.g., short articles, letters, editorials, speeches, brochures, etc.)	68.5
Q35g	Stories or poems	52.1
Q35h	Responses to critiques of written texts	50.3
Q35i	Plays; film; or television scripts	27.3
Q35j	Note taking	74.5

Frequency 1 (5-pt. Scale). Results yielded an overall mean of 3.59 for *Frequency 1*, indicating a general frequency of at least once a week for Q36-Q38. However, the item percentages show less frequent use. The most frequently used student activity is Q38, "organize ideas into a coherent progression of sentences and paragraphs" (76.2% ≥ twice per month, *M*=3.65). On the other hand, the most infrequent activity is Q37, "organize and synthesize information, including consideration of various perspectives" (19.8% ≤ Once per month, *M*=3.52).

Table 9

Category III: Literacy – Frequency 1 (5-pt. Scale)

	Scale Prompt	*N*	*Min*	*Max*	*Mean*	*SD*	*Alpha*
	This school year, how often did you ask students to complete writing tasks in which they must…	**275**	**3.54**	**3.65**	**3.59**	**0.06**	**0.89**
	Item	**Never**	**Once a month or less**	**2-3 times per month**	**1-2 times per week**	**Almost daily**	
Q36	provide evidence to support their ideas and conclusions?	3.7	15.8	25.5	19.1	31.2	
Q37	organize and synthesize information, including consideration of various perspectives?	4.0	15.8	26.8	20.1	26.5	
Q38	organize ideas into a coherent progression of sentences and paragraphs?	3.0	14.1	26.8	17.1	32.2	

Frequency 2 (5-pt. Scale). As indicated in Table 10, the *Frequency 2* set has an overall mean of 3.35, reflecting a general frequency of 2-3 times per month for the typical student activity. The most frequently assigned student activities according to teachers were Q42, "take turns reading aloud" (79.2%, *M*=3.68) and Q41, "read silently" (78.9%, *M*=3.79). The most infrequently given student activities were Q44, "view videos (movies, documentaries) that relate to student learning" (53.0%, *M*=2.54) and Q43, "write multiple drafts" (38.9%, *M*=2.92).

Table 10

Category III: Literacy – Frequency 2 (5-pt. Scale)

	Scale Prompt	*N*	*Min*	*Max*	*Mean*	*SD*	*Alpha*
	How often did you have your students…	**268**	**2.54**	**3.79**	**3.35**	**0.50**	**0.77**
	Item	**Never**	**Once a month or less**	**2-3 times per month**	**1-2 times per week**	**Almost daily**	
Q39	write about something they have read?	4.4	15.1	26.5	14.8	33.9	
Q40	quote or refer to text as evidence for their answers or opinions?	7.7	15.4	25.2	16.1	30.5	
Q41	read silently?	8.1	7.7	23.5	13.4	41.9	
Q42	take turns reading aloud?	6.4	9.1	28.2	16.4	34.6	
Q43	write multiple drafts?	11.4	27.5	24.8	17.4	12.8	
Q44	view videos (movies, documentaries) that relate to student learning?	15.1	37.9	20.1	17.1	4.0	

Frequency 3 (5-pt. Scale). Table 11 reveals the overall mean for *Frequency 3* to be 3.69, indicating that the typical activity is used "Often". The most frequently assigned ("sometimes" or more) student activities in this set were Q48, "use reading comprehension strategies to identify important ideas from text" (86.9%, *M*=3.91), Q51, "define the purpose for reading specific text" (84.9%, *M*=3.88) and Q49, "synthesize information from text while they are reading, not just when they are finished" (84%). The most infrequently used student activities were Q52, highlight the information they need to remember from text" (22.1%, *M*=3.35) and Q50, "skim and scan text to determine how carefully they will need to read" (18.8%, *M*=3.44).

Table 11

Category III: Literacy – Frequency 3 (5-pt. Scale)

	Scale Prompt	*N*	*Min*	*Max*	*Mean*	*SD*	*Alpha*
	This school year, how often did you have your students do the following?	**272**	**3.35**	**3.91**	**3.69**	**0.19**	**0.94**
	Item	**Never**	**Not Very Often**	**Some-times**	**Often**	**Always**	
Q45	use reading comprehension strategies to access content in reading materials (e.g., textbooks, novels).	5.4	7.0	21.5	35.2	25.8	
Q46	create mental images of text as they are reading.	7.0	7.0	21.8	38.3	21.1	
Q47	make inferences (e.g. predictions, conclusions) from text they are reading.	5.0	5.0	17.8	39.9	26.2	
Q48	use reading comprehension strategies to identify important ideas from text.	4.4	4.0	17.4	37.9	31.5	
Q49	synthesize information from text while they are reading, not just when they are finished reading.	4.0	7.0	21.1	36.9	26.2	
Q50	skim and scan text to determine how carefully they will need to read.	6.7	12.1	23.5	36.6	15.4	
Q51	define the purpose for reading specific text.	3.4	6.4	18.8	35.6	30.5	
Q52	highlight the information they need to remember from text.	9.7	12.4	26.2	27.5	19.1	
Q53	synthesize expository text by identifying the most important information from the text.	5.4	7.0	23.5	34.6	24.2	

Frequency 4 (5-pt. Scale). Frequency 4 set has an overall mean of 3.61 (see Table 12). The most frequent teacher activities in this set ("sometimes" or higher) were Q58, "monitor the level of students' comprehension and confusion as they read text" (81.2%, *M*=3.77) and Q57, "model the use of questioning as a reading comprehension strategy for my students" (80.2%, *M*=3.64). The most infrequent teacher activity was Q54 "model reading comprehension strategies that students will be using for the first time" (18.1%, *M*=3.51).

Table 12

Category III: Literacy – Frequency 4 (5-pt. Scale)

	Scale Prompt	*N*	*Min*	*Max*	*Mean*	*SD*	*Alpha*
	This school year how often did you do the following?	276	3.51	3.77	3.61	0.10	0.95
	Item	**Never**	**Not Very Often**	**Some-times**	**Often**	**Always**	
Q54	model reading comprehension strategies that students will be using for the first time.	7.0	11.1	22.1	29.9	23.5	
Q55	provide guided practice to students when they use reading comprehension strategies for the first time.	7.0	8.1	22.8	33.6	21.5	
Q56	give students feedback as they practice reading comprehension strategies independently for the first time.	7.7	8.7	24.2	31.9	19.8	
Q57	model the use of questioning as a reading comprehension strategy for my students.	5.7	6.7	23.5	35.9	20.8	
Q58	monitor the level of students' comprehension and confusion as they read text.	5.0	6.7	18.1	38.6	24.5	

Multiple Response 3 (checklist – 7 choices). Table 13 shows the proportion of respondents who checked each choice in response to Q59, “This school year, which of the following reading comprehension strategies did you teach your students to use?”. The most frequently selected choices were Q59a, “searching for connections between what students know and new information” (85.9%) and Q59c, “drawing inferences” (84.8%). The least frequently selected choice was Q59f, “repairing faulty comprehension” (51.1%).

Table 13

Category III: Literacy – Multiple Response 3 (Checklist)

	Scale Prompt	
	This school year, which of the following reading comprehension strategies did you teach your students to use?	
	Item	**%Checked**
Q59a	searching for connections between what students know and new information.	85.9
Q59b	asking questions of self, author, and text.	69.3
Q59c	drawing inferences.	84.8
Q59d	distinguishing important from less important ideas.	75.2
Q59e	synthesizing information within and across texts.	56.7
Q59f	repairing faulty comprehension.	51.1
Q59g	monitoring adequacy of understanding.	74.8

Agreement 1 (1 item). As shown in Table 14, a majority of respondents agreed that the arts and literacy are used together for teaching and learning in their classroom (80.9%, *M*=3.16).

Table 14

Category III: Literacy – Agreement 1 (4-pt. Scale)

	Scale Prompt	*N*	*Min*	*Max*	*Mean*	*SD*
	To what extent do you agree or disagree with the following:	280	1	4	3.16	0.74
	Item	**Strongly Agree**	**Agree**	**Disagree**	**Strongly Disagree**	
Q60	In my classroom, the arts and literacy are used together as tools for teaching and learning.	31.2	49.7	8.4	3.4	

Category IV: Assessment

Tables 15-16 review items in the Assessment category. Responses in this category, as for most categories, tended to fall at the middle of the scales. The least positive responses related to allowing students to learn self-assessment skills.

Frequency 1 (5-pt. Scale). Table 15 shows that the *Frequency 1* set has an overall mean of 3.09, corresponding to "Sometimes." The most frequent student activity in this set was Q64, "assess their own progress" (79.9%, *M*=3.38). The most infrequent student activity was Q61 "assist you in determining the criteria (e.g., rubrics, scoring guides) for assessing their work" (38.3%, *M*=2.74).

Table 15

Category IV: Assessment – Frequency 1 (5-pt. Scale)

	Scale Prompt	*N*	*Min*	*Max*	*Mean*	*SD*	*Alpha*
	This school year, how often did you have your students do the following?	283	2.74	3.38	3.09	0.31	0.8
	Item	**Never**	**Not Very Often**	**Some-times**	**Often**	**Always**	
Q61	assist you in determining the criteria (e.g., rubrics, scoring guides) for assessing their work.	13.4	24.8	34.2	17.8	5.4	
Q62	use rubrics/scoring guides to assess their own work.	4.7	13.1	30.5	37.9	8.7	
Q63	assist you in identifying goals and objectives for future learning.	8.4	21.1	37.2	24.2	4.0	
Q64	assess their own progress.	2.0	13.1	34.9	35.2	9.7	

Frequency 2 (5-pt. Scale). The *Frequency 2* set has an overall mean of 3.54 (see Table 16), which is somewhat more positive than the *Frequency 1* set. The most frequent teacher activities in this set were Q67, "give feedback to students before moving on in a lesson" (91.3%, *M*=3.82) and Q68, "discuss students' work in progress in terms of its quality" (89.9%, *M*=3.74). The most infrequent teacher activity was Q65, "teach students how to assess their own progress" (17.4%, *M*=3.23).

Table 16

Category IV: Assessment – Frequency 2 (5-pt. Scale)

	Scale Prompt	*N*	*Min*	*Max*	*Mean*	*SD*	*Alpha*
	This school year, how often did you do the following?	284	3.23	3.82	3.54	0.29	0.83
	Item	**Never**	**Not Very Often**	**Some-times**	**Often**	**Always**	
Q65	teach students how to assess their own progress.	5.0	12.4	37.6	34.2	6.0	
Q66	use brief individual conferences to monitor student progress.	4.0	11.1	36.2	34.9	8.7	
Q67	give feedback to students before moving on in a lesson.	0.7	3.4	25.2	48.7	17.4	
Q68	discuss students' work in progress in terms of its quality.	1.3	4.4	27.2	47.7	15.1	

Category V: Leadership

Tables 17-18 review items in the Leadership category. Responses in this category were less positive overall than for other categories. Although moderately positive about the leadership staying abreast of and sharing the goals of the reform plan, the teachers had concerns about of the viability of school and reform administrators intervening in classroom dynamics.

Agreement 1 (4-pt. Scale). As shown in Table 17, the *Agreement 1* set has an overall mean of 2.74, indicating less positive views than the previously discussed sets. The highest level of agreement occurred on Q70, “The principal attends professional development activities for staff, including institutes, workshops, and select coaching sessions” (79.9%, *M*=3.11). There was also general agreement for Q71, “The principal and teachers at this school have a shared vision for student success” (74.5%, *M*=2.94). The highest level of disagreement occurred for Q69, “The principal coaches me on how to deliver effective instruction” (46.6%, *M*=2.45). Q69 has essentially the same level of agreement (45.6%) as disagreement. High levels of disagreement also occurred for Q73,

"Teacher leaders at this school coach me on how to deliver effective instruction" (33.2%, *M*=2.66) and Q74, "Different Ways of Knowing is building my capacity to develop as a leader" (33.2%, *M*=2.64).

Table 17

Category V: Leadership – Agreement 1 (4-pt. Scale)

	Scale Prompt	*N*	*Min*	*Max*	*Mean*	*SD*	*Alpha*
	To what extent do you agree or disagree with the following?	248	2.45	3.11	2.74	0.19	0.88
	Item	**Strongly Agree**	**Agree**	**Disagree**	**Strongly Disagree**		
Q69	The principal coaches me on how to deliver effective instruction.	6.7	38.9	31.9	14.8		
Q70	The principal attends professional development activities for staff, including institutes, workshops, and select coaching sessions.	25.8	54.0	8.1	3.7		
Q71	The principal and teachers at this school have a shared vision for student success.	19.8	54.7	14.4	5.0		
Q72	The principal and teachers at this school share responsibility for data-driven decision making.	14.4	53.4	17.1	6.4		
Q73	Teacher leaders at this school coach me on how to deliver effective instruction.	11.4	48.0	24.2	9.1		
Q74	Different Ways of Knowing is building my capacity to develop as a leader.	11.4	47.7	21.8	11.4		
Q75	Different Ways of Knowing is building my capacity to implement school change.	11.4	49.0	20.1	11.1		
Q76	This school uses a standard protocol for examining student work.	8.1	51.0	24.8	6.0		
Q77	This school uses a standard protocol for examining teacher lessons.	8.1	47.7	23.8	8.1		

Frequency 1 (5-pt. Scale). As shown in Table 18, the *Frequency 1* set had an overall mean of 3.24. The most frequent professional conversation took place, as described in Q81, with "other teachers at this school" (87.2%, *M*=3.77). The most infrequent professional conversations were with (Q78) "the principal" (35.2%, *M*=2.78).

Table 18

Category V: Leadership – Frequency 1 (5-pt. Scale)

Scale Prompt	*N*	*Min*	*Max*	*Mean*	*SD*	*Alpha*
This school year, how often did you engage in reflective, professional conversations about student success with:	281	2.78	3.77	3.24	0.43	0.76

	Item	Never	Not Very Often	Some-times	Often	Always
Q78	the principal?	15.1	20.1	30.5	25.8	2.0
Q79	other administrators at this school?	7.0	16.8	38.3	26.8	4.7
Q80	teacher leaders at this school?	5.0	10.1	30.2	41.3	6.7
Q81	other teachers at this school?	2.0	5.0	23.5	47.0	16.8

Category VI: Implementation

Tables 19 and 20 review items in the Implementation category. Responses in this category were also mixed. Rated more positively were administrative commitment and implementation of curriculum standards. Concerns were expressed in the areas of meaningful and effective teacher involvement in the reform process as well as time constraints.

Agreement 1 (4-pt. Scale). As shown in Table 19, the *Agreement 1* set has an overall mean of 2.91. The highest level of agreement occurred on Q85, "The principal is committed to the success of Different Ways of Knowing at this school" (79.9%, *M*=3.14) and Q87, "The principal participates in meetings with members of leadership teams on a regular basis" (78.9%, *M*=3.04). The highest level of disagreement occurred in response to Q83, "My principal and I have meaningful conversations about his/her visits" (29.9%, *M*=2.72) and Q86, "Teachers are involved in making the important decisions in this school" (28.2%, *M*=2.75).

Table 19

Category VI: Implementation – Agreement 1 (4-pt. Scale)

	Scale Prompt	*N*	*Min*	*Max*	*Mean*	*SD*	*Alpha*
	To what extent do you agree or disagree with the following?	258	2.72	3.14	2.91	0.18	0.9
	Item	**Strongly Agree**	**Agree**	**Disagree**	**Strongly Disagree**		
Q83	My principal and I have meaningful conversations about his/her visits.	11.7	49.3	19.5	10.4		
Q84	The principal seeks feedback from staff and colleagues on a regular basis.	13.8	49.7	16.8	10.4		
Q85	The principal is committed to the success of Different Ways of Knowing at this school.	29.9	50.0	7.0	5.0		
Q86	Teachers are involved in making the important decisions in this school.	16.1	47.7	17.8	10.4		
Q87	The principal participates in meetings with members of leadership teams on a regular basis.	21.8	57.0	7.0	5.4		
Q88	The principal reflects on school performance with staff on a regular basis.	22.1	54.4	10.4	4.7		

Agreement 2 (4-pt. Scale). As shown in Table 20, the *Agreement 2* set has an overall mean of 2.73, indicating more negative views than in other sets in this survey. The highest level of agreement occurred on Q98, "This school uses a common framework that aligns expectations for current student learning with curriculum" (79.2%, *M*=2.97) and Q105, "Different Ways of Knowing can help improve student achievement at this school" (73.8%, *M*=2.92). The highest level of disagreement occurred in response to Q95, "Our school schedule facilitates enough time to support curriculum integration" (46.3%, *M*=2.44) and Q96, "Our school schedule facilitates enough time to support de-tracking" (45.6%, *M*=2.41).

Table 20

Category VI: Implementation – Agreement 2 (4-pt. Scale)

Scale Prompt	*N*	*Min*	*Max*	*Mean*	*SD*	*Alpha*
To what extent do you agree or disagree with the following?	**198**	**2.41**	**3.00**	**2.73**	**0.14**	**0.92**

	Item	Strongly Agree	Agree	Disagree	Strongly Disagree
Q89	Teachers are committed to the success of Different Ways of Knowing at this school.	7.4	49.7	25.8	9.4
Q90	Teachers have a lot of informal opportunities to influence what happens here.	14.4	50.3	21.5	5.7
Q91	I have sufficient materials to develop Different Ways of Knowing lessons.	10.1	47.0	25.5	10.4
Q92	I work with a team of teachers to plan instruction.	6.7	45.6	32.2	8.4
Q93	We follow up on new programs at this school to make sure they are working.	7.0	46.0	28.5	7.4
Q94	Many special programs come and go at this school.	12.8	42.6	28.5	3.7
Q95	Our school schedule facilitates enough time to support curriculum integration.	5.7	39.6	35.6	10.7
Q96	Our school schedule facilitates enough time to support de-tracking.	5.4	32.9	36.6	9.1
Q97	This school uses a common framework that guides curriculum.	11.1	66.4	10.1	1.3
Q98	This school uses a common framework that aligns expectations for current student learning with curriculum.	12.4	66.8	8.7	1.7
Q99	Working conditions at this school support implementation of a common framework that aligns expectations for student learning with curriculum.	10.4	57.4	15.1	6.0
Q100	Working conditions at this school support implementation of a common framework that aligns expectations for student learning with curriculum.	10.7	57.0	16.1	5.4
Q101	Resources at this school are allocated to support implementation of a common framework for curriculum.	10.7	54.0	18.5	7.0
Q102	Resources at this school are allocated to support implementation of a common framework that aligns expectations for student learning with curriculum.	10.7	54.7	16.8	6.4
Q103	The district is committed to success of Different Ways of Knowing at this school.	8.1	51.0	17.8	5.0
Q104	Different Ways of Knowing fits well with the other school improvement efforts at this school.	11.1	53.0	16.8	6.7
Q105	Different Ways of Knowing can help improve student achievement at this school.	14.8	59.1	12.4	4.7
Q106	DWoK is helping this school achieve its goals.	8.7	50.7	19.5	8.1
Q107	DWoK places a burden on this schools resources.	10.7	31.2	38.6	5.0

Differences Between States

An examination of the mean values for each subcategory by state is shown in Table 21. Texas and California had the smallest participation in this survey. California scored somewhat lower in the Standards, Instruction, and Assessment categories, whereas Texas scored lower in the Literacy and Leadership categories. No large differences in responses are shown between any of the states.

Taking each state in turn, relative strengths for California are: Literacy Frequency 2, which asks how often a teacher had students perform enriching activities such as writing about something they have read or referring to text in their answers; Literature Frequency 4 where the teacher was asked how often he/she modeled comprehension strategies or provided guided practice or monitored student comprehension as they read. As a contrast, Texas was weaker in both these areas. These characteristics are dependent upon the amount of classroom time a teacher may devote to enrichment, which may differ between these two states or perhaps be a function of the regional training provided. Texas has a relative advantage in Assessment Frequency 2, which is the amount of time the teacher involved the student in the assessment process and review before going on to the next topic.

Kentucky demonstrated an advantage in Standards Frequency 1, which is the degree to which the teacher incorporated standards and Backwards Planning in classroom dynamics; and in Instruction Percent 1; which looked at the percentage of the lesson time devoted to a standards-based topic in-depth. Finally, Kentucky was more positive with regard to Leadership Agreement 1, Leadership Frequency 1, Implementation Agreement 1 and Implementation Agreement 2. This is an exception to the overall results. Leadership and Implementation were stronger in Kentucky. For New York, relative

strengths were seen in Literacy Frequency 1, Literacy Agreement 1 and Assessment Frequency 1.

Table 21

Means by State

STATE		Standards Agreement 1	Standards Frequency 1	Instruction Percent 1	Instruction Agreement 1	Instruction Frequency 1
California	**Mean**	**2.92**	**3.63**	**4.71**	**3.04**	**3.18**
	N	17	19	20	19	20
	Std. Deviation	0.23	0.46	0.61	0.41	0.55
Kentucky	**Mean**	**3.15**	**3.94**	**5.25**	**3.14**	**3.56**
	N	109	123	119	116	117
	Std. Deviation	0.34	0.67	0.68	0.48	0.71
New York	**Mean**	**3.18**	**3.76**	**5.21**	**3.14**	**3.57**
	N	89	103	98	97	98
	Std. Deviation	0.35	0.68	0.62	0.41	0.63
Texas	**Mean**	**3.14**	**3.69**	**4.80**	**3.17**	**3.34**
	N	49	50	49	48	48
	Std. Deviation	0.45	0.81	0.80	0.42	0.70
Total	**Mean**	**3.14**	**3.82**	**5.12**	**3.14**	**3.50**
	N	264	295	286	280	283
	Std. Deviation	0.36	0.69	0.70	0.44	0.68

STATE		Literacy Frequency 1	Literacy Frequency 2	Literacy Frequency 3	Literacy Frequency 4	Literacy Agreement 1
California	**Mean**	**3.44**	**3.43**	**3.66**	**3.67**	**3.17**
	N	19	19	17	20	18
	Std. Deviation	0.81	0.58	0.56	0.84	0.71
Kentucky	**Mean**	**3.58**	**3.27**	**3.55**	**3.48**	**3.14**
	N	116	112	115	111	116
	Std. Deviation	1.05	0.84	0.87	0.99	0.70
New York	**Mean**	**3.92**	**3.56**	**3.92**	**3.82**	**3.22**
	N	93	93	92	98	100
	Std. Deviation	1.05	0.89	0.93	1.04	0.76
Texas	**Mean**	**3.09**	**3.16**	**3.56**	**3.44**	**3.11**
	N	46	43	48	47	46
	Std. Deviation	1.14	0.80	1.00	1.19	0.82
Total	**Mean**	**3.60**	**3.36**	**3.69**	**3.60**	**3.16**
	N	274	267	272	276	280
	Std. Deviation	1.08	0.84	0.91	1.04	0.74

Table 21 (continued)

Means by State

STATE		Assessment Frequency 1	Assessment Frequency 2	Leadership Agreement 1	Implementation Frequency 1	Implementation Agreement 1	Implementation Agreement 2
California	**Mean**	**2.74**	**3.35**	**2.54**	**3.23**	**2.91**	**2.54**
	N	20	20	18	20	18	11
	Std. Deviation	0.80	0.49	0.47	0.61	0.45	0.41
Kentucky	**Mean**	**3.11**	**3.52**	**2.87**	**3.33**	**3.14**	**2.85**
	N	115	116	102	117	109	85
	Std. Deviation	0.74	0.74	0.54	0.69	0.57	0.41
New York	**Mean**	**3.25**	**3.54**	**2.76**	**3.20**	**2.76**	**2.69**
	N	100	100	88	100	90	69
	Std. Deviation	0.78	0.72	0.55	0.83	0.69	0.53
Texas	**Mean**	**2.85**	**3.64**	**2.44**	**3.09**	**2.65**	**2.58**
	N	48	48	40	44	41	33
	Std. Deviation	0.89	0.75	0.60	0.77	0.74	0.56
Total	**Mean**	**3.09**	**3.54**	**2.74**	**3.24**	**2.91**	**2.73**
	N	283	284	248	281	258	198
	Std. Deviation	0.80	0.72	0.57	0.75	0.66	0.49

Summary and Conclusions

Overall, the teacher survey reflected areas of relative strengths and weaknesses within and across the different implementation categories. Little is known from the present data regarding the contextual variables that may have been present in the separate states and their schools. This factor, along with the large difference in the number of respondents across states, and the inability to identify individual schools from the WestEd data reduce the value of making inter-state comparisons. These caveats notwithstanding, we offer the following general impressions from the survey results.

- Relatively strong implementation areas were in Standards and Instruction revealing teachers' engagement in self-evaluation activities.
- Areas of concern were Leadership and Reform Implementation. The overall implication is that, at many sites, traditional school hierarchies remained intact. Consequently, the teachers' access to modeling of DWoK behaviors (by

administrators or others) or to receive adequate time and resources for implementation was reduced.

- The commitment of administration to the DWoK goals, and the overall fit of the program with the school were strongly supported.
- Results for Kentucky tended to be more positive than those for other states, especially on Leadership and Implementation.

Study II: Coaches' Interviews

Introduction

The purpose of the present study was to assess the perceptions of the instructional and artist-educator coaches as well as the team leader of DWoK in five Kentucky middle schools. Areas assessed included DWoK's design components, effectiveness, implementation quality and progress, and perceived impact.

The present interview data were restricted to the Kentucky schools due to (a) accessibility of key informants, (b) the large number of schools involved (n = 5 out of the 15 in Cohort I), and, most importantly, (c) the availability to the researchers of student-level data for achievement analyses (which was not the case for the other states) to complement the present qualitative results. As previously indicated, this report presents only generic findings to protect the promised confidentiality of the individual schools.

Method

CREP conducted formal structured phone interviews with the instructional coach, artist-educator coach, and team leader at each of the five Kentucky middle schools. The interviews took place on Tuesday, August 26 and lasted approximately one-hour per school. The interviews were conducted with Jo Ann Mosier (Team Leader), Sherri

Beshears-McNeely (Artist-Educator Coach), Lynne Miller (Artist-Educator Coach), Denise Finley (Instructional Coach), and Kathy Lowe (Instructional Coach). Additionally, Dr. Susan Galletti, Vice-President of Middle Grades Research and Development at The Galef Institute, observed all five sets of interviews, added information, and requested clarification as appropriate.

Schools were discussed separately, and interviewees were asked to be as honest as possible with their answers. Responses were then transcribed and synthesized.

Results

The following bulleted statements are the key impressions obtained from the interviews regarding implementation of DWoK at the five Kentucky middle schools.

Implementation Progress

- At three schools, implementation has been going well. At one school, for example, implementation has been "wonderful. . . it is a little bit of heaven." This school was doing well prior to DWoK, and it was reported to continually improve. At another school, implementation has been going well overall, although the school is committed to many other programs and initiatives which has stretched them thin. At another school, implementation is going very well. Implementation was difficult at first, but it became easier and went quickly because the coaches were able to build relationships *through* the work rather than as a *result* of the work or prematurely *before* the work began.
- In contrast, at two schools, DWoK implementation has been either "surface-level" or non-existent. Still, at one of these schools, schedule changes have enabled DWoK to have had a positive impact.

- The coaches worked hard to define and develop the idea that schools were to be "co-developers" in a research project. As it turns out, many schools were interested only in receiving the free professional development services. DWoK emphasizes concepts and processes; the schools, however, thought they would get free "stuff."

Teacher Support

- At three of the schools, there appeared to be solid support for DWoK. At one of these schools, teacher support has been steadily increasing over the past two years, and this was attributed to the support from the principal. Teacher buy-in was very strong at a second school, where the teachers are very happy and friendly. At a third school, teacher support is "phenomenal;" the staff members are highly supportive of DWoK and of each other.
- At two schools, however, teacher support was guarded at best. For example, the staff at one school has not been open to DWoK. The staff members appear to appreciate the coaches' efforts, but they have the attitude, "We'll believe it when we see it." At another school, there is a broad spectrum of teacher-buy-in. Overall, however, teachers have a "Will this last?" attitude. They seem to lack the concentration and the vision needed for success.

Most Effective Elements

- According to the coaches, the most effective element at all five schools was Design Element #3 - "Thinking, viewing, listening, speaking, reading, and writing all day long." Several schools were reported to have a strong commitment to literacy. Additionally, Design Element #1 - "Standards-based curriculum,

instruction, and assessment linked to big idea" – was reported to be a strength at four of the five schools.

Least Effective Elements

- One of the least effective elements, according to the coaches, appeared to be the lack of planning in three of the schools. At these schools, the union has told teachers that they are not required to turn in lesson plans. Thus, planning is inconsistent at best at two of these three schools. The coaches have "met up with a brick wall" here, given that planning impacts all other Design Elements.

Impact on Students

- In at least four of the five schools, DWoK was reported to have positively impacted student sharing and learning. At one of the schools, for example, the students like that the lessons appeal to their learning style. At another school, students do not necessarily know DWoK name per se, but they do reportedly respond well once they are taught the strategies. At a third school, in classes where student inquiry is implemented, students appeared to be happier, and there appeared to be fewer discipline problems. However, there was rarely enough implementation at this school. Finally, at a fourth school, there appears to have been a positive impact on students as a community of learners. The students were reported to be less rigid and less afraid to share. Additionally, students at this school and another school were reportedly more enthusiastic about learning as a result of DWoK.

Impact on Teachers

- There also appears to have been some growth in teacher planning and instruction as a result of DWoK. At one school, for example, teachers have shown growth in

instruction through the year. Specifically, open response assignments were more rigorous and tied to the core units, something that was not seen before. Still, they are still not where the coaches would like them to be. At another school, teacher planning continues to improve.

- The coaches mentioned that in four of the five schools, improvements in teacher sharing and/or engagement in professional development sessions have resulted. At one school, there has been more sharing between teachers and more group work. Teacher sharing continues to be great at another school. At another school, where implementation had been surface-level, this past summer's institute appeared to be successful, in that the teachers were finally on-board with the coaches in terms of buy-in. Similarly, at another school, the teachers have been engaged during professional development sessions.
- In at least two of the five schools, teacher relationships appear to have improved. In one school, for example, there appears to have been some growth in professional relationships between the teachers. At another school, teacher relationships are strong in and of themselves, but DWoK appears to have enabled them to be a better unit of teachers.

Community and Parental Support for DWoK

- Parent support and involvement has been low at most of the schools, and there were reportedly no changes in involvement as a result of DWoK at four of the five schools. At one school, however, there are many DWoK parent volunteers (i.e., volunteers who are specifically assigned to the DWoK program), who read with the students one-on-one or work on displays of student work. Parent involvement at this school was believed to have increased as a result of DWoK.

Study III: Student Achievement, Attendance, and Climate Analyses for Urban Schools

Introduction

The most critical and revealing component of the present evaluation is the analysis of student achievement effects of DWoK for the Middle Grades. For this purpose, a rigorous quasi-experimental design, in which each DWoK school is matched to a similar control school, was employed. The student achievement and attendance study is restricted to the Kentucky cohort based on the availability of test scores, interest by district and school stakeholders, and representative (reasonable) levels of model implementation.

Method

Participants

Due to the inability to randomly assign schools to treatment and control groups, a two-level matching procedure was used to increase the internal validity of the study (Cook & Campbell, 1979). The first level of matching was at the school level and involved checking for similarity in terms of poverty (% of students eligible for participation in the free or reduced-price lunch program), student mobility, attendance, special education (Exceptional Child Education, ECE), and single-parent households.

To further strengthen the similarity of the treatment and control schools, the state's Accountability Index (AI) was also used for this first level matching procedure. The AI is a combination of the academic index and the non-academic index. First, a weighted academic index is calculated for all content area tests administered within a school (e.g., reading, mathematics). Then, after adding weighted non-academic measures (e.g., attendance) and a weighted norm-referenced test, the Comprehensive Test of Basic

Skills (CTBS), the school's AI for a particular year is determined. The weights used for calculating the AI varies for the different content area tests, but it sums to 1.0. The weights used to calculate a school's AI vary slightly depending upon whether the school is an elementary, middle, or high school. The accountability index is considered a high-stakes system because rewards and sanctions are attached to results.

The sample included three treatment middle schools and three matched control schools. The three treatment schools began implementing DWoK in the 2001-02 school year as part of their effort to increase student achievement. Table 22 illustrates the key characteristics by which the treatment and control schools were matched at the baseline year (2000-2001).

Table 22

School Level Matching Data for Treatment and Control Schools at Baseline Year

School	N	% Poverty	% Mobility	% Attendance	% ECE	% Single Household	Accountability Index
DWoK-A	857	47.3	11.3	93.1	14.1	60.6	65.4
Control-A	962	43.2	10.6	94.7	11.0	50.7	64.5
DWoK-B	915	39.1	10.6	93.2	13.1	50.8	58.6
Control-B	625	41.0	13.2	93.1	13.8	53.5	57.0
DWoK-C	945	38.9	12.7	94.4	13.8	54.5	62.3
Control-C	831	43.9	12.0	92.6	14.0	53.5	58.6

Note. Accountability Index includes academic (e.g., reading, math, science, social studies) and non-academic (attendance, retention) indicators.

Only the students who had both CTBS and Kentucky Core Content Test (KCCT) reading scores in grades 6 and 7, respectively, served as the basis for the second level of the aggregated matching procedure (Rossi, Freeman, & Lipsey, 1999), which took place at the student level. The CTBS test was administered to the sixth graders in the baseline

year (2000-2001). The KCCT reading test was administered to the seventh graders in the first program implementation year (2001-2002). Treatment and control students were matched on six variables, namely previous test scores in reading (6^{th} grade CTBS), age, gender, race, free or reduced-price lunch, and single-parent households in the baseline year (2000-2001).

The matching procedure was evaluated using Analysis of Variance (ANOVA). As shown in Table 23, no statistically significant differences were found on age ($F = .01$, $p > .05$), race ($F = .69$, $p > .05$), free or reduced-price lunch ($F = .87$, $p > .05$), and single-parent ($F = 1.86$, $p > .05$) variables. These results confirmed that the aggregated matching procedure was successful in avoiding the need for using statistical controls (i.e., covariates) beyond differences testing baseline year (2000-01) CTBS reading scores ($F = 5.89$, $p < .05$) and gender ($F = 7.61$, $p < .05$).

To determine the magnitude of DWoK-control differences, each mean difference score was divided by the pooled posttest standard deviation. The resulting standardized differences, or effect sizes, provide summaries of the magnitude of each difference in standard deviation units. Note in Table 23 that the effect size for both variables (CTBS scores and female) associated with significant group differences was relatively weak (both *ES*s = +0.14).

Table 23

Baseline Data for the Individual Student Level Analytical Sample (2000-2001)

	DWoK (*n* = 638)		Control (*n* = 647)		
	M	SD	M	SD	*ES*[a]
CTBS Reading pretest	660.68	44.41	655.16	36.94	+0.14*
Age	14.13	0.46	14.13	0.44	0.00
Female	0.51	0.50	0.44	0.50	+0.14*
Minority	0.38	0.49	0.36	0.48	+0.04
Free/reduced lunch	0.39	0.49	0.41	0.49	+0.04
Single-parent home	0.51	0.50	0.48	0.50	+0.06

Note: * $p < .05$; Female = 1, Male = 0; Minority = 1, White = 0; Free/reduced lunch = 1, paid lunch = 0; Single-parent home = 1, Dual-parent home = 0. Original scores for dummy coded variables were percentages.
[a]*ES* = Effect Size (+ *ES* indicates DWoK > control)

Achievement Measures

To compare Program (DWoK vs. control) achievement over time, all available data from the 2000-01, 2001-02, and 2002-03 school years were abstracted from computerized files provided by the school district that served as the research site. The 2000-2001 school year was used as the baseline. Data from the state assessment system (i.e., CTBS, KCCT) and from the school district (i.e., demographics, Stanford Diagnostic Reading Test) were utilized. The CTBS includes reading, language arts, and mathematics subtests. CTBS reading is administered in KY to third, sixth, and ninth graders. The KCCT Reading test, which includes both multiple-choice and open-ended questions, is group administered to fourth, seventh, and tenth graders. The following section describes

the analytical designs and which sets of test scores were examined in the DWoK vs. control group comparisons.

Design and Procedure

A summary of the various analyses employed is provided in Table 24.

Table 24

A summary of analyses comparing DWoK to control schools on student achievement

Outcome Measure	Year[a]	Analysis	Covariate(s)
7th Grade KCCT Reading	2001-02[a]	ANCOVA	6th Grade CTBS Reading Gender
6th Grade CTBS Reading Language Arts Mathematics	2002-03[b]	MANCOVA	Race Poverty[b]
7th Grade KCCT Reading	2002-03[b]	ANCOVA	6th grade SDRT
7th Grade KCCT Humanities & Arts	2002-03[b]	ANCOVA	6th grade SDRT

Note: KCCT = Kentucky Core Content Test, CTBS = Comprehensive Test of Basic Skills, SDRT = Stanford Diagnostic Reading Test

[a]Year 1 of DWoK implementation; [b]Year 2

Year 1 (2001-02). The research design employed a matched treatment-control school pre-posttest design (Cook & Campbell, 1979; Rossi, Freeman, & Lipsey, 1999). Data at the school level for both treatment and control schools were analyzed using descriptive statistics. However to control for prior achievement, data at the student level were analyzed via *ANCOVA*, with the treatment condition as the between-subject factor, the 2000-2001 (sixth-grade) CTBS reading "pretest" scores and gender as covariates, and the seventh-grade KCCT reading scores as the dependent variable (see Table 24). Given

the reasonably large number of students at the participating schools, school pair comparisons were also performed. The purpose of the school pair analysis was to assess differences between each particular treatment school relative to its specific control school.

It should be noted that, for verification purposes, a "full sample" posttest-only analysis was also conducted. For this analysis, all students with 2001-02 KCCT scores were included regardless of the availability of pretest scores. This analysis yielded essentially the same results as the pretest-posttest design.

Year 2 (2002-03). The Year 2 design for sixth grade consisted of a pre-post-test treatment-control group comparison, using *MANCOVA*, on the language arts, reading, and mathematics subtests of the CTBS. The covariate consisted of race and poverty due to the DWoK schools having a greater percentage of minority students and students eligible for free or reduced-price lunch. The Year 2 design for seventh grade consisted of a pre-posttest treatment-control group comparison, using *ANCOVAs*, on the reading and the arts and sciences subtests of the KCCT (see Table 24).

Results

Year 1 Achievement

Overall Sample: 7th Grade KCCT Reading

The *ANCOVA* indicated that the program effect comparing the adjusted KCCT Reading means (M_{adj} = 512.21 and 509.68, respectively) was significant, $F(1,1281) = 3.94$, $p = .047$. $eta^2 = .003$. Computation of the effect size using Cohen's *d* (Cohen, 1988) formula indicated the advantage to be approximately +0.19, as shown in Table 25. The CTBS reading pretest covariate, $F(1,1281) = 1135.73$, $p < .001$, $eta^2 = .470$, and the

gender covariate, $F(1,1281) = 60.05$, $p < .001$, $eta^2 = .045$, were both highly significant. In the latter case, females surpassed males.

Table 25

Unadjusted Means and Standard Deviations for DWoK and Matched-Control Comparisons on KCCT Seventh-Grade Reading Scale Score in 2001-2002 (Year 1).

Comparison Groups	*n*	*M*	*SD*	Effect Size[a]
Overall DWoK Sample	638	514.05	32.26	+0.19
Overall Control Sample	647	507.86	31.28	
DWoK-A School	196	520.94	34.40	+0.18
Control-A School	264	514.77	33.34	
DWoK-B School	210	515.55	33.09	+0.31
Control- B School	164	505.70	29.70	
DWoK-C School	232	506.87	28.03	+0.20
Control-C School	219	501.16	28.15	

[a]Effect sizes were computed from the adjusted means.

School Comparisons: 7th Grade KCCT Reading

Separate inferential tests were performed comparing each DWoK school to its matched control school.

Pair A. The matching procedure was checked for accuracy using independent-sample *t*-tests on the matching variables. No statistically significant differences were found on CTBS pretest scores ($t = 1.09$, $p > .05$), age ($t = 0.20$, $p > .05$), gender ($t = 1.91$,

$p > .05$), race ($t = 1.94$, $p > .05$), free or reduced-price lunch ($t = 0.26$, $p > .05$), and single-parent ($t = 1.78$, $p > .05$). Given the absence of any differences, there was no basis for using any of the demographic variables as a covariate. However, the strong correlation between CTBS reading pretest scores and the dependent variable still justified using *ANCOVA* to increase the power of the group comparison.

As shown in Table 24, the unadjusted DWoK-A mean on KCCT Reading was 520.94 compared to 514.77 for Control-A. The *ANCOVA*, however, yielded a nonsignificant program effect, $F(1,457) = 2.65$, $p = .047$, $eta^2 = .006$. The CTBS Reading pretest covariate, $F(1,1457) = 392.43$, $p < .001$, $eta^2 = .463$, was highly significant.

Pair B. The matching procedure was checked using independent-sample *t*-tests. Statistically significant differences were found only for CTBS pretest scores ($t = 2.43$, $p < .05$) on which DWoK-B ($M = 659.10$) surpassed Control-B ($M = 648.34$). The latter variable was therefore employed as a covariate in the *ANCOVA*.

As shown in Table 24, DWoK-B ($M = 515.55$) scored higher on the posttest than did Control-B ($M = 505.70$). This difference was found to be nonsignificant in the *ANCOVA*, $F(1,371) = 3.16$, $p = .076$, $eta^2 = .008$. The pretest covariate, however, was highly significant, $F(1,371) = 368.84$, $p < .001$, $eta^2 = .499$.

Pair C. Independent-sample *t*-tests, used for verifying the matching procedure, yielded statistically significant differences on race ($t = 4.91$, $p < .05$) only. In the subsequent *ANCOVA*, both race and CTBS pretest scores were used as covariates.

Program means on the KCCT Reading posttest (see Table 24) were directionally higher for DWoK-C ($M = 506.87$) than for Control-C ($M = 501.16$). Differences were significant in the *ANCOVA*, $F(1,447) = 5.56$, $p = .019$, $eta^2 = .012$. Both the race,

$F(1,4477) = 8.09$, $p = .005$, $eta^2 = .018$, and pretest covariates, $F(1,457) = 290.18$, $p < .001$, $eta^2 = .394$, were also significant.

Summary

After one year of implementing DWoK, the three schools demonstrated a significant overall advantage (*ES* = +0.19) over their matched control schools on seventh-grade KCCT Reading. Individual school comparisons lacked comparable power and therefore, although directionally favoring the DWoK schools in each case, were significant only for Pair C. However, the associated effect sizes of +0.18, +0.31, and +0.20 were at least suggestive regarding program impacts. In this regard, Borman, Hewes, Overman, and Brown's (2003) recent meta-analytic study of 29 comprehensive school reform (CSR) models indicated an overall effect size of from +0.10 to +0.14, with the range for the most successful category being +0.17 to +0.21. Only 3 out of the 29 models achieved this high status (Direct Instruction, School Development Program, and Success For All). Clearly, the present first-year findings for DWoK Middle Grades compare favorably to these standards.

Year 2 Achievement

Overall Sample: 6th Grade CTBS Subtests

Analyses of 2002-03 CTBS subtests compared DWoK and control schools on the sixth-grade subtests of reading, language arts, and mathematics. A *MANCOVA* was employed, in which covariates were race ($t = 3.10$, $p = .002$) and free or reduced-price lunch ($t = 2.07$, $p = .04$), which were found to significantly differ between program groups. As an additional check of group similarity, *t*-tests for independent samples were conducted on the Stanford Diagnostic Reading Test and Stanford Diagnostic Math Test

administered at the beginning of the school year. No differences (t = 0.127, p = .90; t = -1.47, p = .14, respectively) were found for either test.

Descriptive unadjusted statistics for each group on the three CTBS subtests are summarized in Table 26. As can be seen, the DWoK mean is directionally higher than the control group mean on all three tests. The *MANCOVA* (Wilks' Lambda) yielded a significant overall program effect, $F(3,1605) = 6.18$, $p < .001$. Follow-up univariate tests were significant (all p's < .001, df = 1, 1607) on all three subtests: language (F = 12.68), reading (F = 15.89), and mathematics (F = 12.77). As indicated in Table 26, the effect sizes associated were relatively small in magnitude ranging from +0.10 to +0.13. Those associated with adjusted means were slightly higher, ranging from +0.16 to +0.18.

Table 26

Unadjusted Means and Standard Deviations for DWoK and Matched-Control Comparisons on CTBS Sixth-Grade Scale Scores in 2002-2003 (Year 2).

Comparison Group and Subtest	*M*	SD	Effect Size[a]
Language			
DWoK	652.64	42.81	+0.11
Control	647.95	41.21	
Reading			
DWoK	654.93	41.11	+0.13
Control	649.78	40.95	
Mathematics			
DWoK	654.09	46.72	+0.10
Control	649.01	50.31	

Note: DWoK n = 825; Control n = 786.
[a]Effect sizes were computed from the adjusted means.

7th Grade KCCT Reading Test Scores

Initial analyses, conducted to verify the similarity of the program groups, failed to show significant differences on any of key demographic variables examined: , age (p = .56), race (p = .12), free or reduced-price lunch status (p = .63), single parent (p = .55), and gender (p = .51). In comparing groups on KCCT reading scores, we conducted two parallel analyses. The first was a posttest-only ANOVA (no covariate) using full-sample data (DWoK n = 934; control n = 909). Results showed a significant Program effect [F(1, 1841) = 13.67, p = .001, ES = +0.17], favoring the DWoK group (M = 512.10, SD = 36.40) over the control group (M = 506.13, SD = 32.76). Unadjusted means and standard deviations relating to this analysis and subsequent comparisons and are presented in Table 27.

The second analysis was an ANCOVA using fall 2001 Stanford Reading Diagnostic Test (SDRT) scores as the covariate. To determine group equivalence on the covariate, we first conducted an ANOVA comparing the 2001 SDRT means for the DWoK and control groups (see Table 27). Results significantly favored DWoK [F(1,1590) = 6.08, p = .014, ES = +0.12]. Adjusting for SDRT effects, the ANCOVA performed on reading scores yielded a significant Program effect [F(1,1589) = 5.81, p = .016, ES = +0.17] favoring DWoK (M_{adj} = 511.78, SD = 35.82) over control (M_{adj} = 508.84, SD = 32.12). The 2001 SDRT covariate was highly significant [F(1, 1589) = 1534.35, p < .001].

7th Grade KCCT Arts & Humanities Test

As in the preceding reading analyses, an ANCOVA using the baseline (2001) SDRT scores as a covariate, was conducted. The dependent variable was the KCCT Arts & Humanities (AH) subtest. . Results indicated that although the SDRT covariate was highly significant [F(1, 1569) = 792.48, p < .001], the program effect did not approach

significance [$F(1, 1569) = .002$, $p = .962$]. As shown in Table 27, the unadjusted KCCT-AH means showed a slight directional advantage for DWoK ($M = 506.41$) over the control group ($M = 500.48$; the adjusted means ($M_{adj} = 503.34$ and 503.50, respectively) were nearly identical.

Table 27

Unadjusted Means and Standard Deviations for DWoK and Matched-Control Comparisons on KCCT and SDRT Covariate Scores for Year2 Analyses

Comparison Group and Test	*M*	*SD*	Effect Size[a]
Fall, 2001 SDRT			
DWoK	661.03	45.64	+0.12
Control	655.79	35.82	
2002-03 KCCT Reading			
DWoK	513.24	35.82	+0.17
Control	507.35	32.12	
2002-03 KCCT AH			
DWoK	506.24	76.02	0.00
Control	500.48	72.43	

Note: DWoK $n = 825$; Control $n = 786$.
[a]Effect sizes were computed from the adjusted means.

Summary

The Year 2 results are generally supportive of DWoK effects. On all three CTBS subtests in sixth grade and on KCCT Reading in seventh-grade, DWoK schools significantly surpassed the matched control schools. No differences, however, were found on the arts and humanities subtest. Although the effect sizes (range = +0.10 to +0.17) for the statistically significant comparisons are modest in size, they still approach

the levels associated with the more successful CSR models identified in the research literature (Borman et al., 2003).

Student Attendance

In 1999, the Kentucky Board of Education adopted the Long-Term Accountability Model. The Accountability Model includes the results of the KCCT and some non-academic indicators such as attendance. For the particular case of the middle school level, attendance has a weight of 3.8% in the Accountability Model. This is a similar weight as the one assigned to the national CTBS (5%).

Overall Sample

Analyses were performed comparing the DWoK and control schools on attendance in the baseline year (2000-01), and Year 1 (2001-02) and Year 2 (2002-03) of DWoK implementation. Means and standard deviations for program and control schools by year are presented in Table 28. Descriptive statistics indicated that attendance rates at DWoK schools decreased by about 0.60 of a percentage point from the baseline year to Year 2 of implementation. During this same period; attendance at the control schools decreased by 1.37 points, over twice the rate of decline for DWoK.

Table 28

Means and Standard Deviations for DWoK and Matched-Control Comparisons on Three-Year Attendance Rates

Comparison Group and DWoK Implementation Year	*M%*	*SD*
2000-01 (Baseline)		
DWoK	95.33	4.18
Control	94.98	4.81
2001-02 (Year 1)		
DWoK	95.06	4.46
Control	94.87	4.89
2002-03 (Year 2)		
DWoK	94.74	5.41
Control	93.61	6.30

The analytical design employed was a repeated measures ANOVA, consisting of a between-subjects factor (DWoK vs. control) by a within-subjects factor (three years of attendance). The attendance rate was computed using the following formula:

$$[(m - a) / m] \times 100$$

In the formula, school membership (m) for the school year minus attendance (a) for the school year is divided by membership (m) for the school year. The resultant dividend is then multiplied by 100 to obtain the attendance rate expressed in percentages.

Findings of interest in the two-way ANOVA primarily concerned the interaction of program x year, which would reflect whether attendance improved in DWoK relative to control schools over time. To a lesser extent, the main effect of program would indicate benefits of DWoK by showing higher attendance in all three years combined (both baseline and post-implementation).

Results indicated a significant program main effect [$F(1, 1245) = 4.68$, $p = .031$, $e^2 = .06$], year main effect [$F(2, 2490) = 43.18$, $p < .001$, $eta^2 = .18$], and program x year interaction [$F(2, 2490) = 9.73$, $p < .001$, $eta^2 = .08$]. Independent samples *t*-tests were used to compare the attendance rates between treatment and control group in each school year. The only significant paired-comparison effect was found between the treatment and comparison group in 2002-03. After two years of DWoK implementation, the DWoK mean ($M = 94.74$) surpassed the control mean ($M = 93.61$), $t(1245) = 3.35$, $p < .001$, *E*S = . +0.19.

School Comparisons

Separate two-way ANOVAs were performed comparing the three-year attendance rates of the individual matched school pairs. As in the analysis for the overall sample, the main interest was in the program x year interaction. Because follow-up pairwise comparisons of program and control means involved multiple independent sample *t*-tests, the Bonferroni procedure (Kirk, 1995, p. 137) was used to control for familywise Type I error rate.

The analysis for Pair A failed to yield a significant program main effect [$F(1, 449) = .45$, $p = .505$] or program x year interaction [$F(2,898) = .58$, $p = .56$]. For Pair B, however, the program x year interaction was significant [$F(2, 712) = 5.16$, $p < .01$], although the program main effect was not [$F(1,356) = 2.234$, $p < .136$]. Follow-up t-tests for Pair B [$t(324) = 2.42$, $p = .016$, $ES = +0.26$] indicated that DWoK ($M = 94.40$, $SD = 5.14$) had a significantly higher attendance rate than the matched control school ($M = 92.94$, $SD = 6.27$) in Year 3. No other significant differences were found.

For Pair C, the analysis yielded a significant program main effect [$F(1, 436) = 11.25$, $p = .001$] and program x year interaction [$F(2, 872) = 18.29$, $p < .001$]. Follow-up

tests showed significant differences in Year 1, $t(118) = 2.44$, $p = .02$, and Year 3, $t(118) = 4.84$, $p = .001$. In Year 1 (baseline), the DWoK school ($M = 95.31$, $SD = 4.27$) had a moderate advantage ($ES = +0.24$) over the control school ($M = 94.24$, $SD = 4.91$). In Year 3, the DWoK advantage increased ($ES = +0.48$) as a result of the control school's attendance ($M = 92.17$, $SD = 6.84$) declining while DWoK's attendance ($M = 95.02$, $SD = 5.41$) remained at a high level.

Summary

Attendance rates at DWoK schools decreased by about 0.60 of a percentage point from the baseline year to Year 2 of implementation. During this same period; attendance at the control schools decreased by 1.37 points, over twice the rate of decline for DWoK. The 2(program) x 3(year) ANOVA showed the interaction pattern to be significant: although there were no program group differences in 2000-01 (baseline) or 2001-02 (Year 1), the DWoK schools significantly surpassed ($p < .001$, $ES = +0.19$) the control schools in 2002-03 (Year 2).

Climate and Perceptual Data

As part of the standard district evaluation process, climate and attitude surveys are administered to teachers, parents, and students at each school. Although procedures for this systemic survey administration do not permit rigorous control over sampling and administration, the return rates are generally judged to include nearly all teachers and over 50% of the students and parents. Thus, results should provide suggestive evidence of the perceptions by major participant groups of school conditions. The following descriptive analyses present the outcomes for each of the DWoK-control school pairs on two of the surveys: (a) School Climate and Atmosphere (20 items), and Quality of Education (25 items). Responses to each item were made on a five-point Likert-type

scale ranging from disagree (1) to agree (5). The School Climate and Atmosphere items deal with respondents' perceptions of pride in the school, expectations of others, positive atmosphere, student motivation, administrative support, and the like. The Quality of Education items solicit evaluations of the academic program, curriculum offerings, Extended School Services, professional development for teachers, and focus on student writing. A summary of overall survey means for the three school pairings in provided on Table 29.

Table 29

Means and Standard Deviations for DWoK and Matched-Control Comparisons on Three-Year Attendance Rates

Matched DWoK-Control Pair and Survey Type				
Pair A		Teacher	Parent	Student
School Climate and Atmosphere				
DWoK	*M*	4.2	4.3	4.0
	(n)	(24)	(124)	(830)
Control	*M*	4.1	3.9	3.7
	(n)	(22)	(202)	(592)
Quality of Education				
DWoK	*M*	4.0	4.3	4.0
	(n)	(24)	(124)	(830)
Control	*M*	4.0	3.8	3.8
	(n)	(22)	(202)	(592)

Table 29 (continued)

Means and Standard Deviations for DWoK and Matched-Control Comparisons on Three-Year Attendance Rates

Pair B		Teacher	Parent	Student
School Climate and Atmosphere				
DWoK	*M*	4.0	3.8	3.6
	(n)	(34)	(161)	(822)
Control	*M*	3.8	3.8	3.7
	(n)	(13)	(207	(507)
Quality of Education				
DWoK	*M*	4.1	3.8	3.8
	(n)	(34)	(161)	(822)
Control	*M*	3.8	3.7	3.7
	(n)	(13)	(207)	(507)
Pair C		Teacher	Parent	Student
School Climate and Atmosphere				
DWoK	*M*	3.5	3.7	3.4
	(n)	(26)	(63)	(751)
Control	*M*	4.0	NA[1]	3.8
	(n)	(24)		(835)
Quality of Education				
DWoK	*M*	3.9	3.8	3.6
	(n)	(26)	(161)	(751)
Control	*M*	4.2	NA[1]	3.9
	(n)	(24)		(835)

[1]Insufficient sample size (< 10)

The results, in general, present a mixed and inconclusive picture of climate and quality of education conditions. Briefly, for Pair A, climate was marginally higher for DWoK relative to the Control school (from 0.1 on Teacher Survey to 0.3 on Parent Survey). On Quality of Education, the only difference favored DWoK by 0.5 on the Parent Survey. For Pair B, the Program means were almost identical on climate, while showing small advantages for DWoK on quality of education. For Pair C, the control school had a noticeable advantage on the Teacher Survey in climate (diff. = 0.5) and to a lesser extent on quality of education (diff. = 0.3). Comparable advantages (0.4 and 0.3, respectively) were indicated for the Control group on the Student Survey.

Taken as a whole, the data do not support the conclusion that conditions at DWoK schools systematically differed from those at Control schools. Results for DWoK-A and DWoK-B were slightly favorable, while a moderate negative pattern was indicated for DWoK-C.

Study IV: Rural Achievement Analyses

Methodology

Sample. The sample included 6th, 7th, and 8th grade students from 2000-2001 to 2002-2003 in two schools: one DWoK school, and one comparison school. The comparison school was selected on the basis of similarity to the DWoK school in terms of the percentage of students eligible for free or reduced-price lunch, student mobility, attendance, special education (Exceptional Child Education, ECE), percentage of single-parent households, and the state's Accountability Index. Three cohorts were present in the sample: 2002 eighth graders, 2003 eighth graders, and 2003 seventh graders. Originally, it was planned to include a second DWoK school in the evaluation, but numerous efforts to find a suitable matched control school were unsuccessful.

Measures. The Reading, Language, and Mathematics subtests of the CTBS9/ Terra Nova were administered to sixth-grade students. Seventh- and eighth-grade students were administered the Kentucky Core Content Test subtests (KCCT) for their respective grade levels. For seventh grade, the KCCT subtests in Reading and Science were administered. For eighth grade, the KCCT subtests in Mathematics, Social Studies, Arts and Humanities, and Practical Living/Vocational Studies were administered. The Arts and Humanities and Practical Living/Vocational Studies subtests were not administered in 2001.

Preliminary data screening. As shown in Table 30, no data were available for 2001 6th graders in the DWoK school, so 2002 seventh-grade cohort analyses were not possible. Two-year match rates for the 2002 8th grade cohort were 183 of 204 cases (89.7%) for the DWoK school, and 177 of 202 cases (87.6%) for the comparison school. Two-year match rates for the 2003 8th grade cohort were 190 of 204 cases (93.1%) for the DWoK school, and 185 of 209 cases (88.5%) for the comparison school. For the 2003 7th grade cohort, two years of matched data were available for 167 of 179 cases (93.3%) for the DWoK school, and 185 of 202 cases (91.6%) for the comparison school.

Table 30

Total Sample Size by Grade, Year, and School[1]

	6th			7th			8th		
School	2001	2002	2003	2001	2002	2003	2001	2002	2003
DWoK	0[2]	370	34	206	206	179	192	204	204
Comp	199	210	191	198	208	202	211	202	209

[1]Maximum effective sample size. Sample sizes for specific analyses are smaller due to missing data for specific tests or lack of matching data in longitudinal analyses. [2] Data not available.

Three 2-way (matched, not matched) analyses of variance were performed on first-year scores to determine whether matching status was likely to bias the results (i.e., whether there was a relationship between pretest and matching status, and whether the relationship varied across schools). Results showed no main effects for any cohort, indicating that student attrition from the study was not related to pretest achievement.

Analyses. The basic model used to analyze achievement results was a multivariate analysis of covariance (MANCOVA), in which the outcome variables were KCCT subtest scale scores for the grade/year cohort, and covariates were scale scores from either the KCCT (eighth-grade cohorts) or Terra Nova (seventh-grade cohorts) administered the previous year. Three MANCOVAs were conducted: 2003 seventh-grade cohort, as well as 2002 and 2003 eighth-grade cohorts. Wilk's Lambda was used as the criterion for multivariate significance. In cases where a significant multivariate effect was indicated, univariate ANCOVAs were performed to ascertain which subtests contributed to the multivariate effect. Post hoc analyses were performed on significant univariate results to determine which means were significantly different. If significant effects were indicated, effect size estimates were computed for each subtest by

subtracting the comparison group mean from the treatment group mean, and dividing by the total standard deviation for that subtest. Descriptive results (average achievement by grade and year) were also tabulated.

Results

Descriptive Profile

For sixth graders, scores on all three subtests were relatively stable in both schools from 2001 to 2003 (see Table 31). Seventh-grade Reading scores remained nearly the same at the DWoK school (from M = 519.9 to M = 520.7), while increasing somewhat at the comparison school from M = 504.0 to M = 513.2 (see Table 32). Likewise, seventh grade Science scores improved somewhat at the comparison school (M= 501.0 to 505.3), while remaining virtually unchanged at the DWoK school (M = 500.9 to 499.6; see Table 32). Eighth grade Mathematics scores remained relatively constant at both the DWoK school (M = 527.7 to 528.3) and the comparison school (M = 528.7 to 531.8; see Table 33). On the Social Studies eighth-grade test, scores remained stable for the DWoK school (M = 514.5 to 512.1), while increasing at the comparison school (M = 502.1 to 511.3; see Table 33).

Table 31

Sixth Grade Terra Nova Mean Normal Curve Equivalent Scores by Subject, Year, and School

	Reading			Language			Mathematics		
School	2001	2002	2003	2001	2002	2003	2001	2002	2003
DWoK	n.c.	52.8	52.4	n.c.	49.4	51.4	n.c.	47.2	45.9
Comp.	49.7	51.2	49.8	46.0	47.2	45.9	50.3	50.6	48.4

Note. Sample sizes are given in Table 1. For normal curve equivalent scores, the norm group mean equals 50, standard deviation equals 21.06.

Table 32

Mean Seventh Grade Kentucky Core Content Scale Scores by Subject, Year, and School

	Reading			Science		
School	2001	2002	2003	2001	2002	2003
DWoK	519.9	515.7	520.7	500.9	502.4	499.6
Comp.	504.0	496.8	513.2	501.0	502.7	505.3

Note. Sample sizes are given in Table 1. Total standard deviations: Reading 2001, 33.4; Reading 2002, 33.3; Reading 2003, 30.90; Science 2001, 30.6; Science 2002, 31.0; Science 2003, 29.1.

Table 33

Mean Eighth Grade Kentucky Core Content Scale Scores by Subject, Year, and School

	Mathematics			Social Studies			Arts & Humanities			Practical Living/Vocational		
School	2001	2002	2003	2001	2002	2003	2001	2002	2003	2001	2002	2003
DWoK	527.7	527.1	528.3	514.5	523.0	512.1	n.a.	522.5	516.3	n.a.	528.3	517.8
Comp.	528.7	521.8	531.8	502.1	515.8	511.3	n.a.	505.4	512.4	n.a.	488.8	503.2
Total SD	36.9	39.4	36.5	45.1	45.7	44.2	n.a.	61.3	77.4	n.a.	60.3	57.6

Note. Sample sizes are given in Table 1. n.a. = test not administered that year.

Inferential Tests of Program Effects

2003 seventh grade cohort. MANCOVA showed a statistically significant treatment effect ($F_{2,346} = 11.06$, $p < .001$). Post hoc analyses were then performed within matched pairs. Follow-up univariate tests revealed significant differences among schools in mean scores on the Reading ($F_{1,352}$=10.93, $p = 0.001$) and Science ($F_{1,352} = 4.86$, $p = 0.03$) subtests. Post hoc comparisons indicated that students attending the DWoK school had significantly higher Reading scores (*M'* = 521.7) than students at the comparison school (*M'* = 514.2) after controlling for prior achievement (*ES* = +0.24; see Figure 1). Students at the comparison school (*M'* = 505.1) had significantly higher scores than students at the DWoK school in Science (*M'* = 500.6, *ES* = -0.15; see Figure 2).

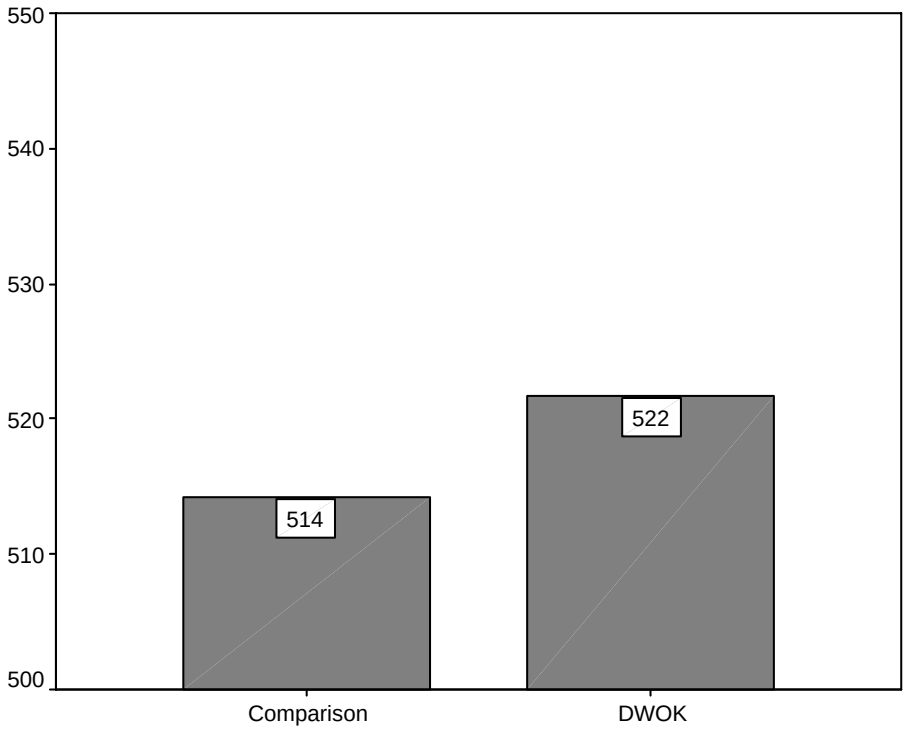

Figure 1. Covariate-adjusted Reading Mean Scale Scores by Treatment, 2003 Seventh Grade Cohort.

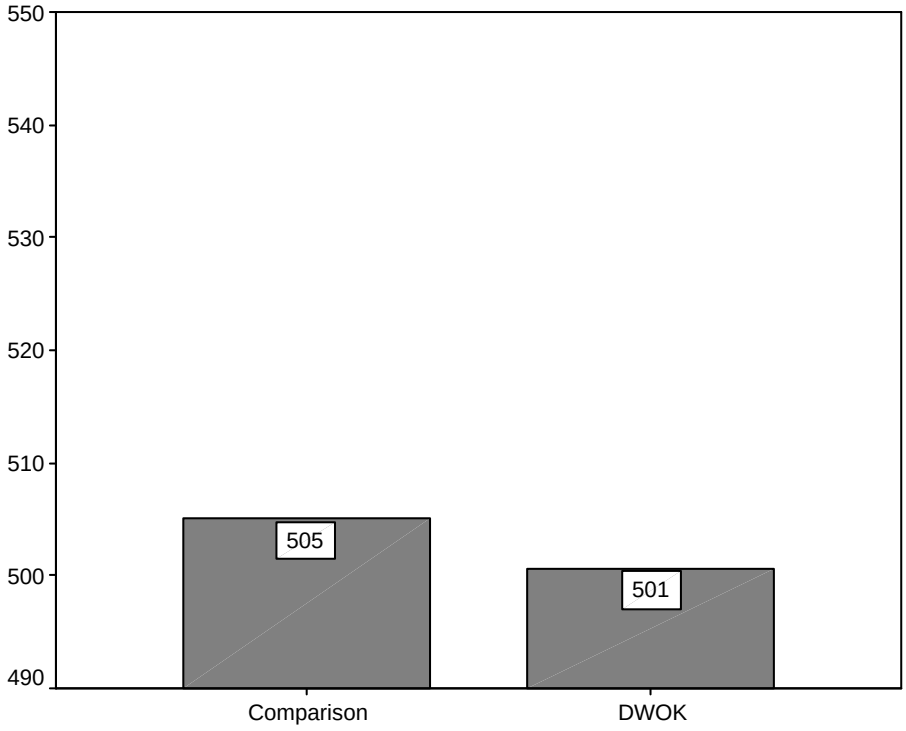

Figure 2. Covariate-adjusted Science Mean Scale Scores by Treatment, 2003 Seventh Grade Cohort.

2002 eighth grade cohort. MANCOVA indicated a significant treatment effect ($F_{4,352} = 8.20$, $p < .001$). Follow-up univariate tests revealed significant differences between schools in mean scores on the Social Studies ($F_{1,355} = 5.24$, $p = 0.023$), and Practical Living/Vocational Studies ($F_{1,355} = 26.24$, $p < .001$) subtests. Post hoc pairwise comparisons indicated the DWoK school had a significantly higher adjusted mean Social Studies score (*M*' = 519.8) than the comparison school (*M*' = 511.7; ES = +0.18; see Figure 3), and a significantly higher adjusted mean Practical Living score (*M*' = 525.2) than the comparison school (*M*' = 499.4; ES = +0.43; see Figure 4). There were no significant differences between the schools in Mathematics (DWoK *M*' = 526.5, comparison *M*' = 526.6) or Arts and Humanities (DWoK *M*' = 520.0, comparison *M*' = 515.6).

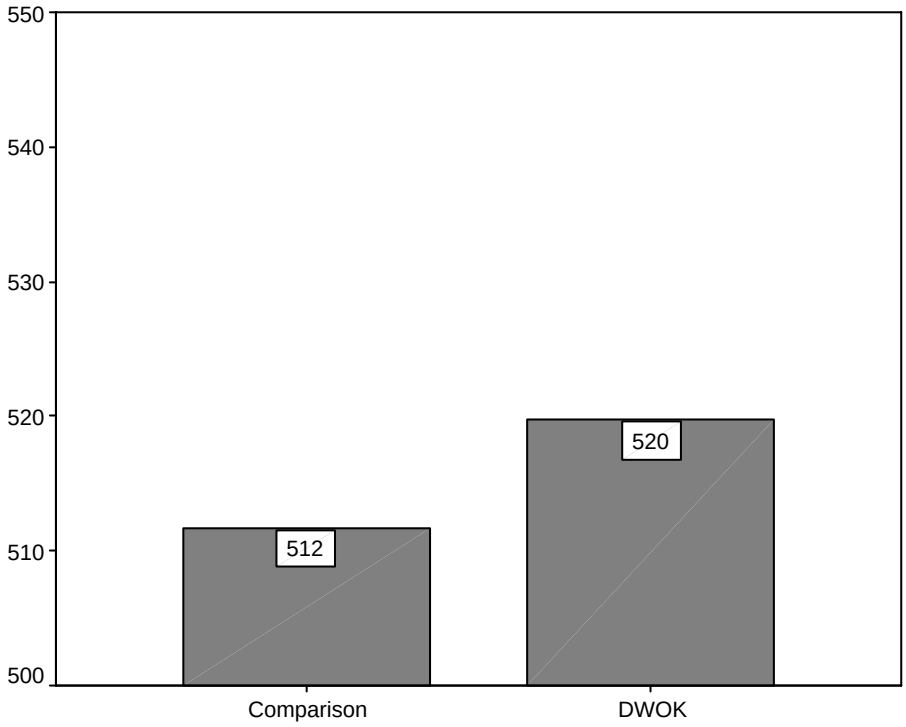

Figure 3. Covariate-adjusted Social Studies Mean Scale Scores by Treatment and School Pair, 2002 Eighth Grade Cohort.

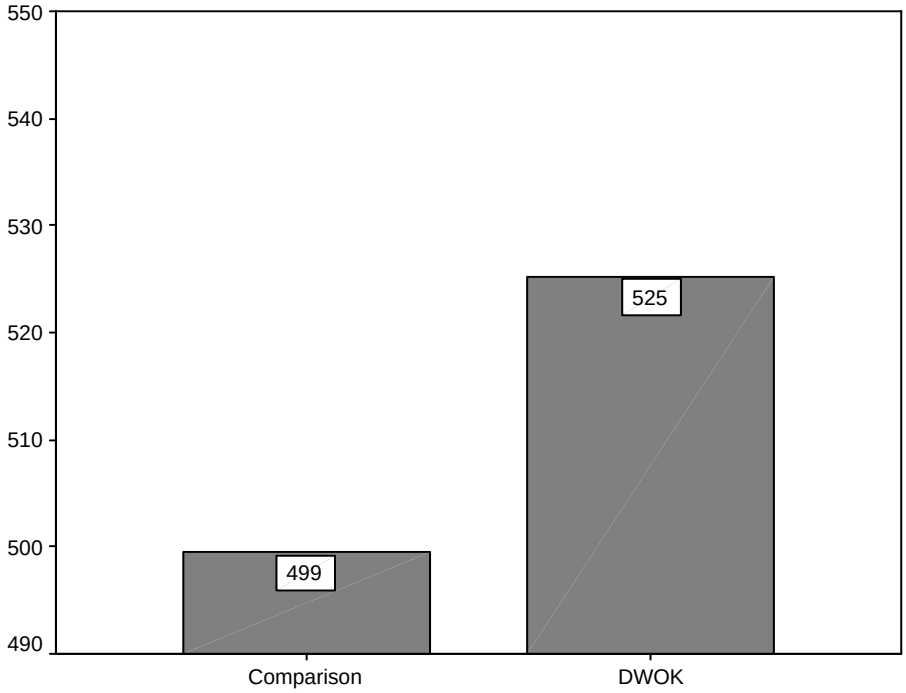

Figure 4. Covariate-adjusted Practical Living Mean Scale Scores by Treatment and School Pair, 2002 Eighth Grade Cohort.

2003 eighth grade cohort. MANCOVA revealed a statistically significant main effect for treatment status ($F_{4,368}$=10.58, $p < .001$). Univariate follow-up tests indicated significant treatment differences on the Mathematics ($F_{1,371}$=22.94, $p < .001$), Social Studies ($F_{1,371}$ = 22.99, $p < .001$), and Arts and Humanities subtests ($F_{1,371}$ = 10.98, p = .001). No effects were indicated on the Practical Living/Vocational Studies subtest (DWoK *M*' = 512.7, comparison *M*' = 512.0). Post hoc tests revealed that, after controlling for prior achievement, students in the comparison school had significantly higher mean scores in Mathematics (*M*' = 537.8; see Figure 5), Social studies (*M*' = 519.2; see Figure 6), and Arts and Humanities (*M*' = 526.9; see Figure 7) than students in the DWoK school (*M*' = 524.7, 505.6, 505.7 for Mathematics, Social Studies, and Arts and Humanities, respectively). Effect size estimates were *ES* = -0.36 for Mathematics, *ES* = -0.31 for Social Studies, and *ES* = -0.27 for Arts and Humanities.

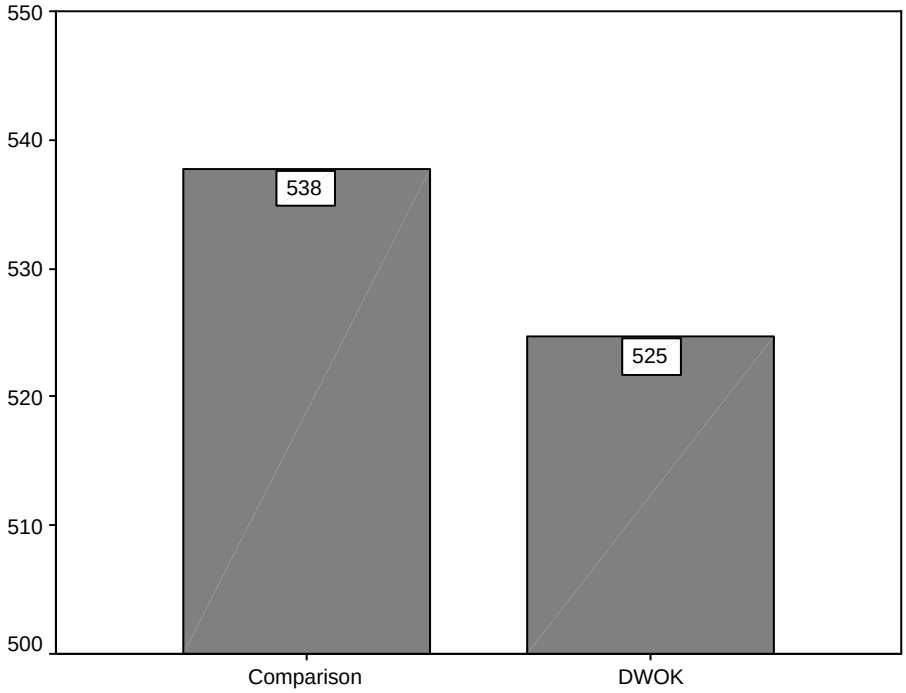

Figure 5. Covariate-adjusted Mathematics Mean Scale Scores by Treatment, 2003 Eighth Grade Cohort.

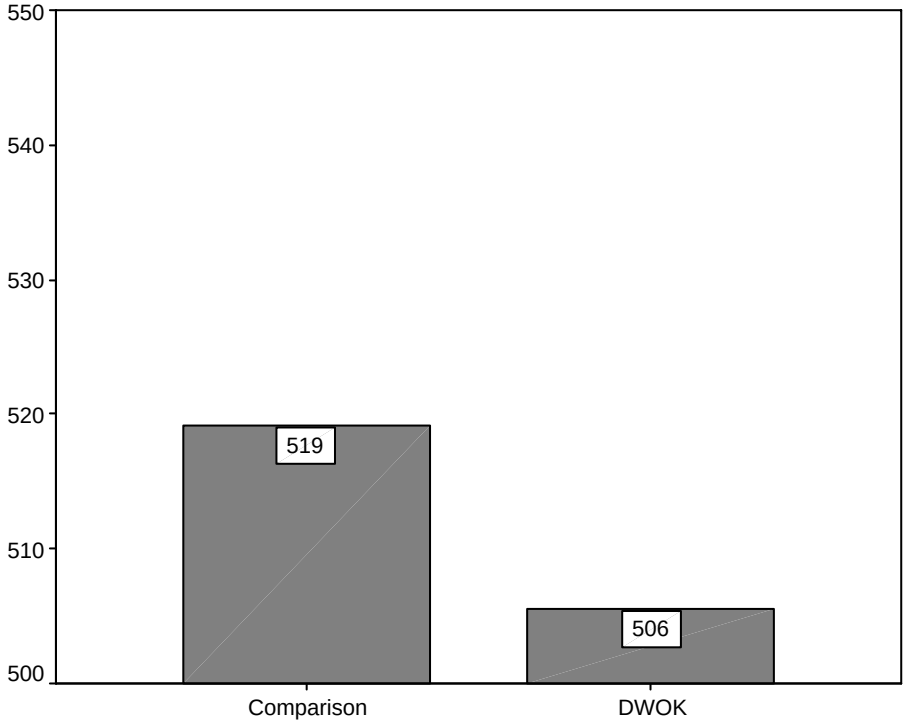

Figure 6. Covariate-adjusted Social Studies Mean Scale Scores by Treatment, 2003 Eighth Grade Cohort.

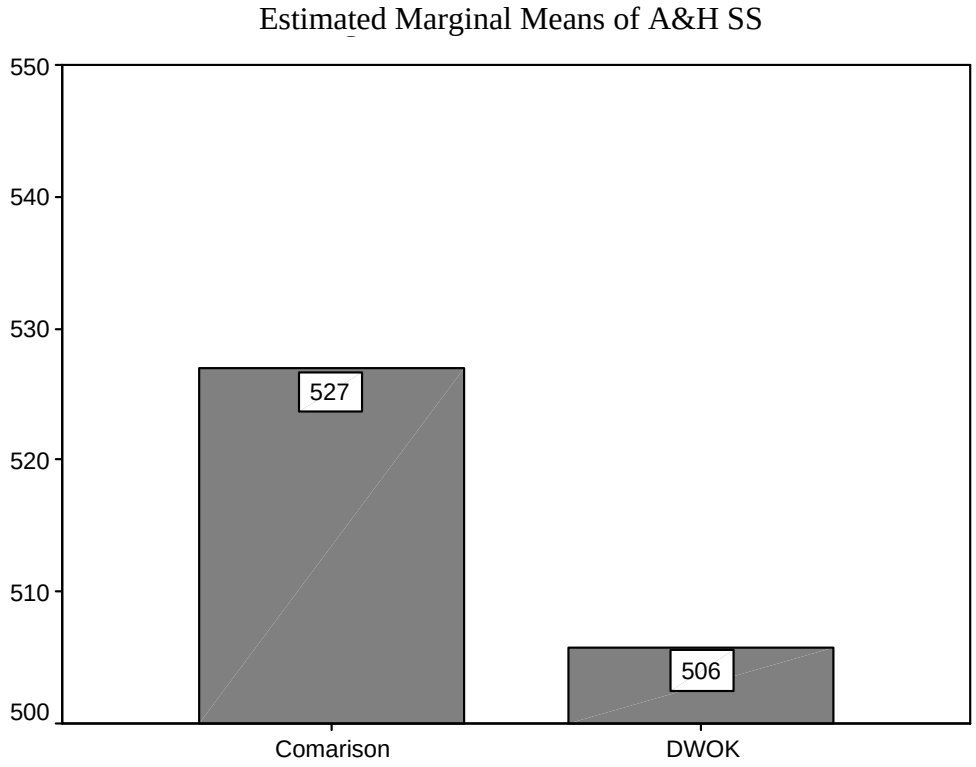

Figure 7. Covariate-adjusted Arts and Humanities Mean Scale Scores by Treatment, 2003 Eighth Grade Cohort.

Findings and Discussion

Inferential tests of program effects showed a mixed pattern of results that generally favored the comparison school. For the 2003 7th grade cohort, the comparison school had a statistically significantly higher mean in Science, although the effect was small (*ES* = -0.15), while the DWoK school had a significantly higher mean in Reading (*ES* = +0.24). Results for the 2002 8th grade cohort favored the DWoK school in Social Studies (*ES* = +0.18) and Practical Living (*ES* = +0.43). For the 2003 8th grade cohort, results clearly favored the comparison school. The comparison school had significantly higher mean scores in Mathematics (*ES* = -0.36), Social Studies (*ES* = -0.31), and Arts and Humanities (*ES* = -0.27).

Discussion

This section will summarize the findings from all four sub-studies as they relate to each of the three research questions.

1. What is the quality of the *Different Ways of Knowing* model?

Strengths. One strength, according to the teachers, appeared to be the emphasis on standards. Almost all teachers agreed that their lessons have specific learning goals, that they analyze student work to modify their instruction, that they provide continuous feedback to their students while they are working on projects, and that all the learning goals of the lessons are based on standards.

Another strength appeared to be perceptions of Instruction. Teachers most commonly agreed that they plan lessons that provide opportunities for continuous learning, build on students' knowledge and skill through research and collaboration, and develop students' expertise, deep understanding, and presentation skills. Teachers were also positive about Assessment, with most agreeing that they provide student feedback prior to moving on in a lesson and discuss the quality of students' work in progress.

Weaknesses. Teachers appeared to most frequently agree about their lack of confidence in their school administration and DWoK staff to provide them with the time, resources, and expertise to improve their own performance as well as the schools' performance. Although moderately positive about the leadership staying aware of and sharing the goals of the reform plan, the teachers had concerns about the administration intervening in classroom dynamics. For example, less than half (45.6%) agreed that either the principal or teacher leaders coach them on how to deliver effective instruction. About 60% agreed that Different Ways of Knowing is building their capacity to develop as a leader or implement school change.

2. How well is the model being implemented at the school sites? High/low fidelity?

Implementation strengths based on the teacher survey centered on the commitment and involvement of the school administration to DWoK goals, alignment with curriculum standards, and the overall fit of DWoK with their school. Implementation concerns were expressed in the areas of meaningful and effective teacher involvement in the reform process as well as time constraints in order effectively support curriculum integration and de-tracking.

Teachers appeared to be more positive about DWoK implementation at the Kentucky schools than the other states. Still, implementation at the Kentucky schools appears to have differed widely depending on the school. At three schools, implementation has been going well. At one school, for example, implementation has been "wonderful." This school was doing well prior to DWoK, and it was reported to continually improve. At another school, implementation has been going well overall, although the school is committed to many other programs and initiatives which has stretched them thin. At another school, implementation is going very well. Implementation was difficult at first, but it became easier and went quickly because the coaches were able to build relationships *through* the work rather than as a *result* of the work or prematurely *before* the work began.

At two schools, however, DWoK implementation has been either "surface-level" or non-existent. At one of these schools, teachers have a "Will this last?" attitude, and they seem to lack the concentration and the vision needed for success. At the other school, teachers take what they want to incorporate, and the coaches may not be told about it. The staff appear to appreciate the coaches' efforts, but they appear to want the

proof that it works first, rather than being on board to try it. Additionally, at both schools, the union has told the teachers that they do not have to plan lessons. At one of these schools, however, there appears to have been a positive impact due to schedule changes.

3. In what ways has the model resulted in improvements in student achievement, teacher practice, and other components of comprehensive school reform?

DWoK appears to have resulted in improvements in several areas of comprehensive school reform. In the three urban Kentucky schools, DWoK appears to have had a positive impact on student achievement, as compared to matched control schools. After the first year of DWoK implementation, the three schools demonstrated a significant overall advantage over their matched control schools on seventh-grade KCCT Reading. Given a lack of power, individual school comparisons were significant only for Pair C. Furthermore, on all three CTBS subtests (i.e., math, language arts, and reading) and in KCCT Reading in Year 2, DWoK schools surpassed the matched control schools. Although the effect sizes were modest in size, they still approached levels associated with more established CSR models (Borman et al., 2003). Of further note, attendance patterns were significantly more positive over the three-year study period (from baseline year to Year 2) in DWoK than in control schools.

Results at the one rural school, however, were equivocal. Overall, 2003 outcomes favored the DWoK school in Reading in seventh-grade, but the comparison school on several of the KCCT subtests in eighth grade. In 2002, the eighth-grade results favored DWoK on multiple tests.

According to the coaches, there appear to have been positive effects of DWoK on student sharing, learning, and enthusiasm. In at least four of the five schools, DWoK

appears to have positively impacted student sharing and learning. Since implementation at one of the schools, for example, there has been more sharing among students. At another school, students do not necessarily know DWoK per se, but they do, according to the coaches, respond well once they are taught the strategies. At a third school, in classes where student inquiry is implemented, students appeared to be happier, and there appeared to be fewer discipline problems. However, there was rarely enough implementation at this school. Finally, at a fourth school, there appears to have been a positive impact on students as a community of learners. The students were reported to be less rigid and less afraid to share. Additionally, DWoK students at this school and another school were reportedly more enthusiastic than prior to the program.

Finally, according to the coaches, there appear to have been positive effects of DWoK on teacher instruction, planning, sharing, and engagement. In at least four of the five schools, there have been improvements in teacher sharing and/or engagement. Progress was also reported in teacher instruction and planning. At one school, for example, open response assignments, for example, were more rigorous and tied to the core units, something that was not seen before. Finally, in at least two of the five schools, teacher relationships appear to have improved or remained strong.

References

Borman, G. D., Hewes, G., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research, 73*(2), 125-230.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally.

Herman, R. (1999). *An educators' guide to schoolwide reform.* Arlington, VA: Educational Research Service.

Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (1999). *Evaluation: A systematic approach* (6th ed.). Newbury Park, CA: Sage.

Appendix A. Teacher Survey Results Summary[2]

Standards.	
Teachers most commonly agreed that: • their lessons have specific learning goals (98.6%). • they analyze student work to modify their instruction (98.4%). • they provide continuous feedback to their students while they are working on projects (94.6%). • all the learning goals of the lessons are based on standards (94.3%).	Teachers least commonly agreed that they: • have available time to plan standards-based instruction with colleagues (39.9%). • provide students with the opportunity to critique each other's work (77.8%). • always introduce the standards used in their lessons to students (78.2%).
The most frequently used strategy was aligning classroom assessments to standards (96%).	The least frequently used strategy was using backward planning to develop lessons (82.6%).

Instruction	
Teachers most commonly agreed that they plan lessons that: • provide opportunities for continuous learning (96%). • build on students' knowledge and skill through research and collaboration (89.6%). • develop students' expertise, deep understanding, and presentation skills (89.2%).	Teachers least commonly agreed that they plan lessons: • based on what students already know (72.5%).
The most frequently used instructional activities were: • explaining to students how a current lesson is linked to prior lessons and future lessons (94.3%). • providing students with examples of exemplary work to clarify expectations (91.3%).	The least frequently used instructional activities were: • the Backwards Planning Template (66.4%). • the use of guiding questions with the Questioning Triangle to relate the "big idea" to lesson content (68.5%).

Literacy	
The most frequently indicated:	The least frequently indicated:
student reading materials were: • textbook chapters (71.5%). • non-fiction books or articles (70.4%).	student reading materials were • vocational, work-related, or other functional print materials (39.1%). • novels (41.9%).
student writing activities were: • note-taking (74.5%). • worksheet exercises (71%).	student writing activities were: • plays, film, or television scripts (27.3%).

[2] For agreement items, in this table, both positive and negative items are referenced based on the percent agreeing or strongly agreeing. Thus, in contrast to the text, negative items are presented in reference to the *least frequently agreed* items (as opposed to the *most frequently disagreed* items). In the text, "Agreement" refers to a positive response and "Disagreement" to a negative response.

Similarly, for frequency items, negative items are presented in reference to the *least frequently used* items (as opposed to the *most infrequently used* items). Percentages were based on those indicating the item was used at least sometimes (i.e., sometimes, often, or always) as compared to rarely or not at all.

student reading and thinking tasks were: • taking turns reading aloud (79.2%). • reading silently (78.8%).	student reading and thinking tasks were: • viewing videos that relate to student learning (41.2%). • writing multiple drafts (55%).
student comprehension strategies were: • using reading comprehension strategies to identify important ideas from text (86.8%). • defining the purpose for reading specific text (84.9%). • synthesizing information from text while reading, not just when they are finished (84.2%).	student comprehension strategies were: • highlighting the information they need to remember from text (72.8%). • skimming and scanning text to determine how carefully they will need to read (75.5%).
teacher modeling strategies were: • monitoring the level of students' comprehension and confusion as they read text (81.2%). • modeling the use of questioning as a reading comprehension strategy (80.2%).	teacher modeling strategy was: • modeling reading comprehension strategies that students will be using for the first time (75.5%).
comprehension strategies taught this year were: • searching for connections between what students know and new information drawing inferences (85.9%).	comprehension strategy taught this year was: • repairing faulty comprehension (51.1%).
A majority of respondents agreed that the arts and literacy are used together for teaching and learning in their classroom (80.9%).	

Assessment	
The most frequently indicated student assessment activity was: • students assessing their own progress (79.8%).	The least frequently indicated student assessment activity was: • students assisting teachers in determining the criteria for assessing their work (57.4%).
The most frequently indicated teacher assessment activities were: • providing student feedback before moving on in a lesson (91.3%). • discussing students' work in progress in terms of its quality (90%).	The least frequently indicated teacher assessment activity was: • teaching students how to assess their own progress (77.8%).

Leadership	
Teachers most commonly agreed that: • the principal attends professional development activities for staff (79.8%). • the principal and teachers have a shared vision for student success (74.5%).	Teachers least commonly agreed that: • the principal (45.6%) or teacher leaders (59.4%) coach them on how to deliver effective instruction. • Different Ways of Knowing is building their capacity to develop as a leader (59.1%).

Implementation	
Concerning the administration, teachers most commonly agreed that: • the principal is committed to the success of Different Ways of Knowing at this school (79.9%). • the principal participates in meetings with members of leadership teams on a regular basis (78.8%).	Concerning the administration, teachers least commonly agreed that: • they have meaningful conversations with the principal about his/her visits (61%). • the principal seeks feedback from staff and colleagues on a regular basis (63.5%).
Concerning the school and community, teachers most commonly agreed that: • their school uses a common framework that guides the curriculum (77.5%) and that aligns expectations for current student learning with the curriculum (79.2%). • Different Ways of Knowing can help improve student achievement at their school (73.9%).	Concerning the school and community, teachers least commonly agreed that: • their school schedule facilitates enough time to support: o curriculum integration (45.3%). o de-tracking (38.3%). • DWoK places a burden on their school's resources (41.9%)