

Using Data to Guide and Inform School Improvement

Laura Hassler Lang, Gail Ogawa, Dawn M. Ossont,
Cheryl Nahmias, Alysia D. Roehrig, Christine E. Johnson and Bahia Diefenbach,
with Shana Goldwyn and Amy Albee-Levine
The Learning Systems Institute, Florida State University

Paper presented at the annual meeting of the
American Educational Research Association
San Francisco, CA.

April 7 - 11, 2006

This research was funded by the Multi-University Reading, Mathematics and Science Initiative (MURMSI) from a grant awarded by the U.S. Department of Education to the Learning Systems Institute, Office of the Provost, Florida State University (FY 04 award number U215K040242).

Abstract

In order to increase the effectiveness of reading instruction, educators must have the knowledge and skills to use reading assessment data in meaningful ways to guide decision making. The purposes of this study are to determine 1) educators' baseline knowledge, skills and attitudes related to using student performance data in reading to guide school reform and instructional decision making and 2) the effectiveness of an online course in addressing identified gaps in knowledge, skills and/or attitudes required to use data for decision making. Initial results from the three pre-test measures indicated that principals, teachers, and reading coaches in the randomly assigned treatment and control groups lacked essential knowledge and skills required to use data effectively. Following the successful completion of an online course in applied data analysis, the treatment group scored significantly higher than the control group on basic assessment knowledge and skills related to reading. The results provide a better understanding of the difficulties educators face in turning student assessment data into useful tools for improved teaching and learning. Further, they suggest coursework that might be incorporated in principal and teacher initial preparation and professional development programs to address this issue.

Using Data to Guide and Inform School Improvement

In recent years, educational leaders have faced intensified demands to employ data-driven decision making when implementing school-based instructional programs (Institute for Educational Leadership, 2000; Johnson, 2002; Massell, 2000). School and district officials have increasingly experienced external pressures from federal- and state-mandated accountability systems to provide data-based evidence for student learning outcomes (Murphy & Beck, 1994). While there is some disagreement regarding the efficacy of these accountability systems, many researchers and practitioners have recognized their potential for creating schools that better meet the instructional needs of all students, especially those traditionally short-changed by our schools: children of color and those from high poverty homes (Fuller & Johnson, 2001; Scheurich & Skrla, 2003; Valencia et al., 2001). Yet, while measures of student performance are critical to determining the extent to which issues of equity and social justice are addressed, the use of disaggregated data to examine differences among disparate student groups is a relatively new concept (Johnson, 2002). The *No Child Left Behind Act of 2001* (Pub. L. No. 107-110) places a renewed emphasis on student achievement results, with an added caveat: the requirement that all sub-groups make adequate yearly progress and ultimately achieve high performance standards. The National Education Goals Panel (2000) notes that schools can achieve these high standards by “us[ing] performance information to determine where they were succeeding and where they need to direct their efforts for improvement” (Rothman, 2000, i). The Panel identified “using data to drive improvement” as an important condition necessary for all students to learn.

Not surprisingly, practices associated with the systematic use of data to monitor student progress and evaluate school performance are supported in the effective schools literature

(Levine & Lezotte, 1990; Sammons, Hillman, & Mortimer, 1995). Lezotte (2001) found, for instance, that in schools where all students experience increased achievement, regardless of socioeconomic status or family background, the results of assessments are used to improve teaching and learning at the individual student level and to improve the curriculum as a whole. Skrla, Scheurich & Johnson's (2000) research in schools that effectively address issues related to equity found that these schools utilize assessments of student learning with greater frequency than the required annual outcome measures mandated by states. Finally, in a recent meta-analysis of quantitative research examining the effects of leadership on student achievement, Waters, Marzano and McNulty (2003) identified "monitor[ing] the effectiveness of school practices and their impact on student learning" (p. 4) as a leadership responsibility with a statistically significant relationship to student achievement.

Defining data-driven decision making

Data-driven (or data-based) decision making involves gathering data to determine whether a class, school, or district is meeting its purpose and vision, which then enables the selection of appropriate interventions to meet specific goals. With regard to the most central goal of schools—student learning—the call for data-driven decision making is based on the assumption that collecting and analyzing data to inform instructional decisions helps to elucidate the causal link between particular school policies or teaching practices and student performance. Subsequently, the most successful educational interventions can be utilized to produce optimal student achievement (Miller, 2000). Drawing on systemic reform theory and knowledge management theory, Mason (2003) provides the following definition for effective data use: "Learning from data means to transform data into information as it is interpreted in context. Data

then becomes knowledge as it is shared, applied, and used to promote change and improvement throughout an organization” (pp. 8-9).

Data-driven decision making in education derives from the continuous improvement principle central to Deming’s concept of Total Quality Management (Sagor & Barnett, 1994). Deming’s Plan-Do-Study-Act (PDSA) cycle is commonly reflected in the stages of collaborative action research—problem formulation, data collection, analysis, reporting, and modifying practices (Sagor, 1993). With its emphasis on data-based decision making, action research provides a systematic inquiry framework that is more effective than the more informal or intuitive methods often employed by teachers when evaluating their own practice (Blasé & Blasé, 2000; Delong & Wideman, 1999; Calhoun, 1994). Likewise, at the school-wide level, organizations that fail to analyze and discuss data “are unlikely to identify and solve problems that need attention, identify appropriate interventions to solve those problems, or know how they are progressing toward achievement of their goals. Data are the fuel to reform” (Killion and Bellamy, 2000, p. 1).

Obstacles to Effective Data Use

Despite the increased emphasis on using assessment results to target students’ individual instructional needs and drive school-level reform, teachers and principals have faced difficulties in turning these data into tools for improved teaching and learning (Massell, 2000; Chen, Salahuddin, Horsch & Wagner, 2000; Hassler, Buck & Torgesen, 2004). For example, instructional leaders often review data sources more to satisfy administrative requirements than to assess the quality of teaching and learning in the school (Creighton, 2001b) and school districts sometimes generate more sources of data than school-based personnel can effectively use (Wallace, 1996). Yet the ability of school staff to gather and interpret data is central to the

school's capacity to engage in data-based decision making. According to Ackley (2001), this process goes far beyond the ability to read the scores that students receive on various assessments. Rather, it requires sophisticated interpretations of student performance data to guide instruction, to make informed curricular decisions, and to monitor the effect these decisions have on student performance.

Stiggins (2001) defines an "assessment-literate faculty" as one that can both gather dependable information about student learning and use that data to effectively maximize student achievement. In their proposed fail-safe schools framework, Bellamy, Crawford, Marshall & Coulter (2005) cite periodic assessment of student learning and the effective use of these results to amend instruction as critical characteristics of schools that could be classified as high reliability organizations. For example, they emphasize staff skills as a limiting factor in the use of data for the early detection of students who are struggling. Earlier research findings also suggest that there are gaps between the training that educators received in assessment and the level of assessment literacy they require in order to effectively monitor student learning (Stiggins & Conklin, 1988; Shafer & Lissitz, 1987; Airasian & Jones, 1993).

More recently, Mason (2003, p. 13) analyzed teachers' data use in six Milwaukee schools and found that among the skills that teachers felt they needed in order to use data effectively were assessment literacy; alignment among assessment, curriculum, instruction and standards; technology (computers, software, databases); and data skills (management, analysis, and application). Further, those same teachers recommended professional development and administrative support as two of the resources necessary to learn these skills and to be more effective in their data use.

Role of the Principal

A leader who understands data-based decision-making is a critical component in facilitating teachers' ability to use data effectively (Thornton and Perreault, 2002). In other words, building capacity for teachers to use data effectively is a necessary, but insufficient, component of successful data-based decision-making in schools. The role of the principal as instructional or transformational leader is critical in this process (Blasé & Blasé, 2000; Hallinger, 2000; Sammons et. al, 1995). For example, Fox (2001) emphasizes the role of the principal in making sure that teachers use classroom assessment data in their daily instructional decision making, and he suggests that leaders must do so by modeling data-based decision making themselves.

Indeed, authors of texts used in principal preparation programs agree that principals need to become more adept at evaluating test data (Drake & Roe, 2003; Popham, 2000). According to Hess and Kelly (2005), while no less than 63% of superintendents report that raising student achievement is the biggest part of principals' evaluations, an analysis of 210 course syllabi from 56 education-administration preparation programs reveals that only 16% of all class sessions addressed issues related to "managing for results." McNamara (2000) argues that in order to successfully develop the required skills, practicing principals need to engage in professional development that stresses, among other things, data analysis skills and illustration of statistical concepts using school-related, real-world data. Furthermore, Pounder and Crow (2005) recommend involving teachers with leadership aspirations in collecting, analyzing and reporting typical school data that addresses issues related to accountability and equity.

Online Professional Development for Educational Leaders

Given the perceived gaps between the training that educators receive in assessment and the level of assessment literacy they require (Airasian & Jones, 1993; Shafer & Lissitz, 1987; Stiggins & Conklin, 1988) the question then becomes how to best train practitioners to effectively generate and analyze data to improve their students' learning. Districts and schools are looking increasingly toward alternative delivery models for professional development. While many continue to debate the effectiveness of distance education (DE) versus face-to-face professional development, it is a rapidly expanding medium for delivery of professional development for educators. Promising features of DE include increased access to a broad array of universities, greater flexibility, cost savings, and greater opportunity for collaboration with other educators (Killion & Bellamy, 2000).

In their meta-analysis of professional development research, Hawley and Valli (1999) recognized "sustained engagement" (as opposed to one-time, face-to face models, for example) as one of the most important design principles for effective professional development. Given the expense and scheduling obstacles associated with a sustained face-to-face model, online models provide a viable alternative in meeting this important goal of effective professional development (Anderson & Henderson, 2004).

In a recent meta-analysis of the comparative distance education research from 1985 to 2002, researchers (Bernard, et al., 2004) found that on measures of student achievement, attitude, and retention for online versus classroom-based courses, pedagogy tended to take precedence over the media for course delivery, particularly with regard to achievement outcomes. Furthermore, research into the characteristics of online learners indicates that certain demographic variables like age and gender may serve as markers for characteristics related to

success in an online course. In studies of distance learners in North American higher education, the distance education students tend to be mostly female and older than their on-campus university counterparts (Thompson, 1998). Based on their analysis of online courses offered in the 64 colleges of the State University of New York (SUNY), Fredrickson and Picket (2000) offer this explanation for success in online instruction:

Often, older students—especially those with familial obligations—are seeking further education out of necessity.... They tend to have higher expectations, more motivation, and a more serious attitude for a number of reasons. If the courses are well designed, it is not unreasonable to expect these students to participate at higher levels and to experience higher levels of satisfaction and learning because of their backgrounds. (p. 20)

Given that most K-12 teachers are within this demographic, it is reasonable to expect that an online course would be a successful professional development option for that population.

Research Focus

While relying on student performance data to guide classroom and school-level decision making appears to be grounded in theory and research, many teachers and principals are unable to use the practice effectively. A number of scholars in the field identify less than adequate preparation in what has been termed “assessment literacy” (Bracey, 2000) as an underlying cause. In order to systematically explore this assertion, this study sought to 1) establish baseline measures of educators’ knowledge and perceived skills about using student performance data to guide decision making and 2) determine whether an online course is effective in addressing identified gaps in knowledge, skills and/or attitudes required to use data for decision making.

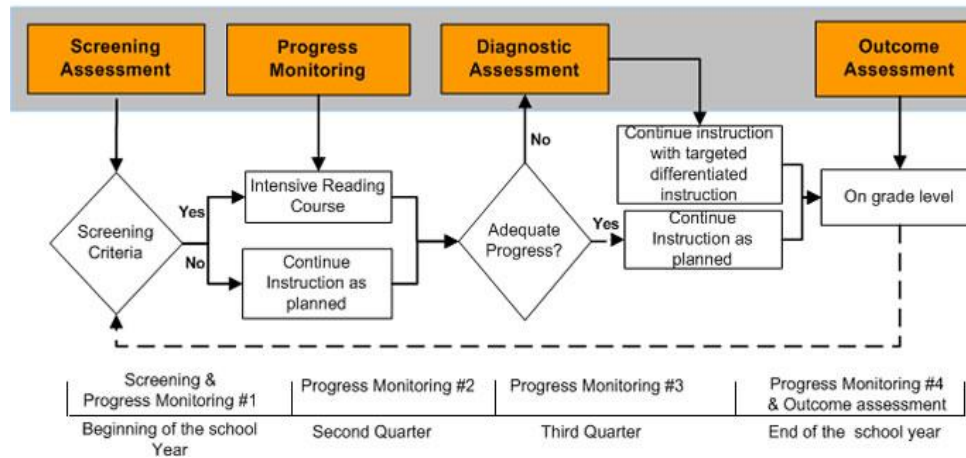
Reading assessment. The first step in undertaking the study was to identify a student performance domain that was commonly assessed and important in practical terms. Reading was

selected for several reasons. Nowhere is the increased emphasis on data more keenly felt than in the realm of reading achievement. The *No Child Left Behind (NCLB) Act* requires by the 2005-2006 school year that states test annually in reading and math in grades 3-8 and once in high school. According to Education Week's *Quality Counts* report, for the 2004-05 school year, 46 states have standards-based tests in place in reading and math at the elementary, middle, and high school levels (Skinner, 2005). While some states test various other domains (science, social studies, writing), reading and mathematics serve as the centerpiece for most accountability systems. For educators and policy-makers, the focus on reading derives from both an understanding of the impact that reading ability has on achievement across content areas and an awareness of the increased literacy demands that students face in their personal lives as citizens and workers (Moore, Bean, Birdshaw & Rycik, 1999; Snow, 2002).

Assessment model. While it is beyond the scope of this paper to fully describe school-level assessment practices, the reading assessment model (Fig. 1) utilized in this study includes screening, progress monitoring, diagnostic and outcome measures of student performance. It is aligned with federal requirements for assessment of student performance as described in the *No Child Left Behind Act of 2001* (Pub. L. No. 107-110). Related concepts were used in defining the knowledge, skills and attitudes required for principals and teachers. Additionally, principal and teacher standards, and competencies related to student assessment in general, and to reading more specifically (e. g., Florida Educator Accomplished Practices (Florida Education Standards Commission, 1996); Florida Reading Endorsement Competencies), were used in the process of identifying the required knowledge base and skills. Examples of knowledge and skills include, for example, knowledge of reading components and reading assessment, interpretation of results,

applications for multiple forms of tests and measures, skills in action research, and determining reliability and validity.

Figure 1. Reading Assessment Flowchart (Hassler & Ogawa, 2002)



Principals, teachers, and reading coaches each play critical roles in implementing school- and classroom-level plans based on this model. Teachers use initial screening measures to identify students' initial instructional needs and to target those students who might require additional attention in order to be successful. Progress monitoring provides teachers with intermittent information regarding the extent to which students are on track to achieve annual learning outcomes and to change instruction for those students who are not making incremental gains. The results of progress monitoring provide critical information for principals and coaches to pinpoint teachers whom might need professional development, such as in-class modeling, and to identify weaknesses in the curriculum or in the implementation of specific programs.

When students fail to make adequate progress after adjustments are made in instruction, diagnostic testing provides teachers with more information about the individual student's specific strengths and weaknesses. Outcome measures, typically given in the fourth quarter, provide information about students' annual learning gains. These norm-referenced and criterion-

referenced tests, typically identified as the high-stakes assessments, are used in measuring the extent to which the school is successful in meeting the overall instructional needs of students. Principals provide these results to teachers to identify their relative strengths and weaknesses, and to the school community to use in setting annual learning goals. Outcome measures from the prior year are also used in the initial screening of students the following year. In that roles of principals, teachers, and reading coaches are inter-dependent, examining them in isolation would not address important aspects of using student data to drive decision making. This approach is aligned with our extant understanding of both the direct and indirect (including teacher-mediated) effects of leadership on student performance and the extent to which these variables impact student performance (Leithwood, Louis, Anderson & Wahlstrom, 2004).

Given the wide variations in assessment practices from state-to-state, district-to-district, and even school-to-school, identifying a sample of sufficient size using common reading assessment measures required that the study include a number of schools within a given state administering common outcome measures of reading performance.

Methodology

Applied Data Analysis for Principals and Teachers (ADAPT) was a randomized controlled trial designed to determine the impact of data-based instructional and curricular decision making on middle school student performance in reading, as measured by two state-wide outcome assessments, the Florida Comprehensive Assessment Test of the Sunshine State Standards in Reading (FCAT SSS-Reading), a criterion-referenced measure, and the Florida Comprehensive Assessment Test- Norm-Referenced Test (FCAT-NRT Reading). In undertaking the investigation, ADAPT Phase I involved the determination of educators' baseline knowledge,

skills and attitudes related to using student performance data to guide decision making. Further, it investigated the extent to which an online course was effective in addressing identified gaps in required knowledge, skills and/or attitudes. ADAPT Phase II was undertaken as a second study to explore the extent to which teachers, reading coaches and principals actually used reading assessment data to inform instructional and curricular decisions after the online course ended. ADAPT Phase III, planned as a two-year follow-up investigation to explore the impact of data-driven decision making on student performance in reading, will be completed in 2006. The third phase is described in more detail in the Discussion section of this paper.

Methodology –Phase I

Phase I was designed to answer the following research question:

Is an online course on reading assessment effective in addressing identified gaps in the knowledge and skills required to use data for decision making?

To ensure that the principals, reading coaches, and teachers in the treatment group for this study possessed the requisite knowledge and skills to analyze data, participants were enrolled in a sixteen-week, 3-credit graduate-level course online professional development course entitled Applied Data Analysis for Educational Leaders, EDA 5931 offered through the Florida State University College of Education during the 2004 fall semester. The development, implementation and outcome model for the entire ADAPT project is presented in Figure 2.

Participants

The research project was conducted in 25 middle schools in Florida, recruited via a statewide conference for middle school principals and an email invitation from Florida's Education Commissioner. Principals, in turn, recruited assistant principals, reading coaches and teachers at their respective schools. To be eligible to participate, each middle school was required to have a

reading coach, core and supplemental reading programs, and a reading assessment plan in 2004-05. In addition, voluntary consent to participate in the research project had to be obtained from the principal (or assistant principal), reading coach, and at least three language arts or reading teachers from any grade level who did not teach the exact same group of students.

A total of 26 public middle schools volunteered and met eligibility criteria. These schools were matched in pairs based on combined z-scores including SES (percent eligible for free/reduced price lunch) and percent of 8th graders scoring at level 1 or 2 on the 2004 FCAT-SSS in reading. One school from each pair was randomly assigned to the treatment group (n=13); the other to the control group (n=13). After random assignment, one treatment group school withdrew from the project. Descriptive data for the 25 remaining schools are shown in Table 1. Participating schools were located throughout the State of Florida. Eligibility for free or reduced price lunch varied from 20% to 78%. Racial/ethnic composition of the student population ranged from only 3% Black and Hispanic at Thomas Sims Middle School to 84% Black at Memorial Middle School and 80% Hispanic at Miami Springs Middle School. In all but one school, more than 30% of 8th graders scored below state standards (levels 1 or 2) in reading on the criterion-referenced portion of the Florida Comprehensive Assessment Test – Sunshine State Standards (FCAT-SSS). Approximately half of the schools had a higher percentage of students performing below state standards than the 2004 state average¹ (56% of 8th graders at levels 1 or 2) on this measure. In the 25 participating schools, 152 individuals voluntarily consented to participate. Three of the treatment group teachers withdrew before the course started. Table 2 summarizes the total number of participants for each group at the start of the study. Participants

¹ <http://fcat.fldoe.org/2004/default.asp> - see *Reading Scores: Statewide Comparison for 2001 to 2004*

TABLE 1

Demographic data for participating schools

School	District	FL region	Low SES ² (percent)	Black (percent)	Hispanic (percent)	Low reading performance ³ (percent)	School grade ⁴
Treatment schools							
Arvida	Dade	South	34	13	58	48	B
Belle Vue	Leon	North	74	65	3	75	C
Davidson	Okaloosa	North	26	8	4	42	A
Forest Grove	St. Lucie	South	73	40	19	67	C
Fort Clarke	Alachua	North	38	32	6	46	A
FSU School	FSU	North	21	25	8	36	A
Highlands	Duval	North	51	53	2	71	C
Oak View	Alachua	North	61	30	2	63	B
Osceola	Pinellas	Central	52	28	4	51	A
Memorial	Orange	Central	62	84	10	84	D
South Miami	Dade	South	50	25	56	50	A
Westwood	Alachua	North	47	34	7	49	B
Control schools							
Howard Bishop	Alachua	North	61	54	3	51	A
Kenwood K-8	Dade	South	41	4	63	42	A
Key Largo	Monroe	South	44	7	26	50	A
Largo Middle	Pinellas	Central	47	15	7	58	B
Miami Springs	Dade	South	74	12	80	71	C
Nautilus	Dade	South	78	7	72	61	B
St. Cloud	Osceola	Central	38	2	12	61	B
Southwood	Dade	South	31	26	32	37	A
Osceola	Okeechobee	South	52	7	20	59	A
Lake Stevens	Dade	South	77	46	49	73	C
Lincoln	Alachua	North	58	59	2	48	B
Thomas L. Sims	Santa Rosa	North	20	2	1	23	A
Yearling	Okeechobee	South	61	6	9	69	B

² Source: Florida School Indicators Reports, <http://info.doe.state.fl.us/fsir/>. SES and race/ethnicity data are reported for 2002-03. Low SES is measured by the percentage of students eligible for free and reduced price lunch.

³ Source: <http://fcats.fldoe.org/2004/default.asp>. Low reading performance is defined as scoring at level 1 or 2 in reading on FCAT-SSS March 2004. Data are reported for 8th graders.

⁴ Source: <http://schoolgrades.fldoe.org/0405/>. Includes 2004 school grades awarded by the Florida Department of Education based on student performance on the Florida Comprehensive Assessment Test.

in the treatment group completed special student applications to Florida State University unless they were already enrolled as students. After their admission, they were registered for the online course, which was limited to study participants only.

TABLE 2

Study Participants, by Experimental Condition and Position

Position\Condition	Treatment schools (n-12)	Control schools (n-13)	Total
Principals	11	11	22
Assistant Principals	3	3	6
Reading Coaches	13	12	25
Teachers	43	48	91
Other*	4	0	4
TOTAL	74	74	148

* Includes 2 learning specialists, 1 staff development specialist, and 1 district-level reading coach

Incentives

For successful completion of the course (grade of B or better), the treatment group participants were offered \$400, three graduate course credits at Florida State University, and state credit for fulfilling Florida Reading Endorsement Competency 3. The stipend (\$400), tuition (\$620) and textbook (\$35) fees were paid by a grant from the U.S. Department of Education. The control group participants were offered \$200, also grant-funded, for completion of the same pre- and post- measurement instruments administered to the treatment group.

Online course content and design

The online course, Applied Data Analysis and Assessment for Educational Leaders, was developed by the FSU Learning Systems Institute in partnership with the North East Florida

Educational Consortium (NEFEC). The course was modeled after face-to-face training workshops for school principals previously offered by the principal investigator and one of the course instructors.

Briefly, the course was designed to provide participants with a foundation in assessment, analysis of student performance data, and the instructional components of reading. It includes an action research project completed by each student to demonstrate the application of these skills. Completion of the course meets state requirements for Florida Reading Endorsement Competency 3 (Foundations of Assessment), one of six competencies that must be fulfilled by Florida teachers to earn K-12 reading endorsement in an add-on certification program. The course was also aligned with Florida's Educator Accomplished Practices (Florida Education Standards Commission, 2003) and the Florida Professional Development System Evaluation Protocol Standards (Florida Department of Education, 2002).

The primary authors for the online course content were four faculty members at the FSU Learning Systems Institute with doctorates in instructional systems design, educational research and evaluation, educational leadership and educational psychology. The reading content was provided by a Ph.D. faculty member at the FSU Florida Center for Reading Research and a doctoral student in middle and secondary education with a specialization in adolescent literacy. Many of the application activities were developed in partnership with principals and teachers who developed and participated in a pilot project during the 2003-2004 school year. The course was delivered using the Blackboard 6 (Bb6) e-Learning System (Blackboard Inc., 2004). The course textbook was Albert Oosterhof's (2003) *Developing and Using Classroom Assessment* (3rd ed.).

The course syllabus (see Appendix A) describes the course goals and objectives. Briefly, these include (1) acquiring the knowledge and skills to analyze and interpret data, (2) identifying appropriate reading strategies, (3) using an assessment framework to guide planning for assessment and curriculum revision, and (4) applying action research tools to reading assessment data to improve instruction and student performance in reading. The course consists of nine instructional units, each containing between two and five lessons. Units are organized based on components of the reading assessment framework: screening assessment, progress monitoring, diagnostic assessment, and outcome assessment. The four components of the assessment framework are presented in the sequence they would normally occur during a school year (see Figure 1).

All of the lessons emphasize primary reading components specified by the National Reading Panel (2000): phonological and phonemic awareness, phonics, fluency, vocabulary, and comprehension. These five components are introduced in a lesson at the beginning of the course and are integrated into subsequent lessons throughout the course. Content related knowledge and skills are reinforced through lesson activities, which include practice applications, and three action research practice activities. The interactive learning devices of Bb6 such as the discussion board, e-mail, and World Wide Web functions are also integrated into the course.⁵ Each lesson can be printed from a separate PDF file. Students use Microsoft Excel (2003) for applications that require data analysis and graphing and are provided with an optional Excel tutorial through the Bb6 system. Internal links present additional examples, information, or guidelines pertinent to each lesson. For example, students can click on key vocabulary terms in the lessons to access definitions in the course glossary. Other internal links produce “pop-ups” with concrete

⁵ For more information on the Blackboard system go to: <http://www.blackboard.com>

examples of instructional strategies mentioned in the lessons. Students have access to a complete reference list and a list of resources (websites and publications) organized by unit and by lesson.

External links provide students access to tools and resources on the World Wide Web, such as the National Reading Panel's recommendations for teaching vocabulary, Just Read, Florida's secondary reading assessment framework and a list of valid and reliable reading assessments compiled by the Florida Center for Reading Research (FCRR).

Course instruction.

The treatment group completed the online course from August, 2004 through February, 2005. All participants were required to have access to a computer, Internet connection and an Internet browser version of 4.0 or higher. Students were able to access the course via the World Wide Web 24 hours a day, seven days a week.

The course was taught by three faculty instructors and five mentors. The three instructors were three of the primary authors of the course (see *Online Course Content and Design*). Four of the mentors were teachers hired by the Leon County, Florida school district and certified in middle school language arts by the National Board for Professional Teaching Standards (NBPTS). The fifth mentor was a doctoral student in middle and secondary education and a member of the instructional design team. During summer 2004, mentors successfully completed the online course and mentor training offered by the FSU Office for Distributed and Distance Learning (ODDL), for which they were paid \$1,500.

Treatment group participants were assigned to one of the six class sections each containing an average of 11 students taught by an instructor-mentor team. All of the treatment group principals and assistant principals were purposefully assigned to the same section. Teachers and reading coaches were randomly assigned to one of the remaining five sections.

The mentors were the primary instructional contacts for students in their respective sections. They initiated and responded to email communications, graded student assignments, monitored student progress, provided instructional feedback, and answered most of the questions raised, contacting instructors when assistance or further expertise was needed. The instructors supervised the mentors, explained course content and provided students additional instruction as needed. They revised course material based on mentor input and monitored the progress of unit tests and practice activities. A separate discussion board within the Blackboard system provided a forum for mentors and instructors to share materials such as email templates and discuss developments in the course, including general problems, content, and instructional questions.

All six sections were taught using the same content and online instruction. At the end of each lesson, students were directed to a second course site where they could complete lesson activities, take tests, monitor their grades, participate in discussion boards, send email, and read course announcements. The unit tests contained between 15 and 30 selected-response items, which were automatically scored and recorded by the Bb6 system. Unit tests and practice applications at the end of each lesson provided opportunities for practical application and immediate feedback on the material taught. Some lesson applications were created by the course development team (such as having students create scatter plots in Excel using their own students' reading achievement data); others were available through a website that supplemented the course textbook (Oosterhof, 2003). Mentors manually graded practice assignments using scoring rubrics and recorded grades in the Bb6 grade book. After receiving feedback from mentors, students were allowed to re-submit assignments for additional points.

During the course, students were required to submit three action research practice activities in preparation for the final action research project plan due at course end. Mentors used

scoring guides to provide formative feedback and grade these activities as complete/incomplete. The action research final project plan was graded by mentors and reviewed by at least two instructors using a scoring guide. The project posttest (see research instruments) served as the course final examination. Student grades were calculated based on the weighted system outlined in the course syllabus.

The State of Florida paid for the first 47 hours of mentoring during the fall 2004 online course as part of the 96 mentor hours required for an annual salary bonus paid to national board certified teachers in Florida. Mentors were paid \$2,000 for additional mentoring hours beyond the initial 47 hours. Course policy dictated that mentors and instructors respond to student e-mail enquiries within 24-48 hours, but generally responses occurred in 24 hours or less. In addition, mentors and instructors were available online 4 hours per week for regularly scheduled office hours. Technical assistance to all of the students and instructors was available 24 hours a day through the Office for Distributed and Distance Learning (ODDL) at Florida State University.

While the course was self-paced, completion of unit tests and activities were required by a due date specified in the course schedule, which was adjusted as necessary to accommodate students affected by a tropical storm and three hurricanes that occurred during August and September 2004. The number of days closed for the participating schools ranged from 2 to 28. However, there was no significant difference ($t = -.23, > .05$) between the average numbers of days the treatment (mean = 6.08) and the control (mean = 6.63) group schools were closed due to the storms.

For research purposes, the conclusion or cut-off date for participation in the research study was February 10, 2005⁶. At that time there were 57 students from the treatment group and 68 participants from the control group eligible to contribute data to the study.

Research instruments

The effectiveness of the on-line course was measured by three instruments administered to both the treatment and control groups: The Test of Assessment Skills and Knowledge (TASK) test, the Survey of Reading Endorsement Competencies (SREC) and the Survey of Concerns Related to Reading Assessment Training (SCRRT). These instruments were designed to measure pre-post treatment changes in student knowledge, skills and attitudes about reading assessment. Instruments were administered at the beginning and end of the course. In addition, an end-of-course evaluation survey was administered to the treatment group only.

Test of Assessment Skills and Knowledge (TASK). A table of specifications outlined the content, cognitive level, and format (multiple-choice and matching) of the items and guided the development of the 99-item test of knowledge and skills. The table was the test blueprint and the method for ensuring the content validity of the test. The course text (Oosterhof, 2003) was a source for most of the test items. The test also included items that were specific to reading and reading assessment. Table 2 shows the general content of the test grouped into 17 categories and the number of items per content category. Reading and measurement experts reviewed the test items, and the instrument was tested with a group of teachers (n = 19) during a 2004 pilot project. The alpha coefficient for the pilot test was 0.77. The alpha coefficient for the TASK pretest was 0.78 and for the posttest was .91 as calculated using Cronbach's alpha.

⁶ The remainder of the students who had not completed the course dropped the course voluntarily or their course registrations were administratively cancelled at the end of Spring Semester 2005 without a grade penalty.

TABLE 2

Test Content and Number of Items

Content • Focus	# of Items	Content • Focus	# of Items
Action Research	1	Reading Components	7
Cognitive Ability the Item Measures	4	<ul style="list-style-type: none"> • Phonics • Phonemic awareness • Fluency • Vocabulary • Comprehension 	
Criterion & Norm-Referenced Interpretations - Frames of Reference	6	Reliability	11
Formal And Informal Tests	4	<ul style="list-style-type: none"> • Method of determination • Generalizability • SEM 	
Formative/Summative Purposes - Test Data	3	Special Groups: LEP & ESE Students - Assessment Accommodations	4
Grading Procedures	2	Test Formats	10
Measurement Terms	2	<ul style="list-style-type: none"> • Traditional and alternative • Scoring options 	10
Performance Objectives	2	Tools for Analysis	
Quantitative/Qualitative Data	3	<ul style="list-style-type: none"> • DART Model • Scatter Plots • Disaggregated data • Expectancy tables • Item analysis 	13
Reading Assessment	9	Types of Test Scores	
<ul style="list-style-type: none"> • Screening • Progress Monitoring • Diagnostic Assessment • Outcome Assessment 		<ul style="list-style-type: none"> • Standardized test scores • Normal Curve 	
		Validity – Types Of Evidence	8

The Survey of Reading Endorsement Competencies (SREC). The measure of assessment foundations as outlined by the FREC-3 competencies was the 18-item self-assessment survey. The participants rated their level of competence for each of the assessment competency questions on the 5-point Likert-type scale ranging from 1 (*expert*) to 5 (*novice*). The alpha coefficient of the scale was 0.96 in the 2004 pilot project. The reliability estimate for this instrument was high as the alpha coefficient for both the SREC pre and posttest was 0.95.

The Survey of Concerns Related to Reading Assessment Training (SCRRTAT). A 35-item self-assessment instrument was adapted from a Concerns Based Adoption Model (CBAM) questionnaire (Hall, George and Rutherford, 1979/1998). The conceptual basis of the CBAM is the assumption that people's concern about an innovation change developmentally. The authors of the CBAM instrument report test/retest reliabilities ranging from 0.65 to 0.86 and alpha coefficients ranging from 0.64 to 0.83 (Hall et al. p.11). Similarly, to the CBAM questionnaire, in the SCRRTAT participants rated their level of concern (how they felt) about being involved with reading assessment. The rating scale contained 7 points grouped within four categories defined as 0 = irrelevant to me now, 1-2 = not true of me now, 3-5 = somewhat true of me now, and 6-7 = very true of me now. The SCRRTAT reliability index (alpha) was 0.82 for the pretest and 0.85 for the posttest.

Course Evaluation Survey. The treatment group students completed the 36-item Course Evaluation Survey (CES) online after they finished the course. The CES contained 36 items. Students rated their level of agreement from 1 (*strongly agree*) through 5 (*strongly disagree*) on 25 statements related to current and future use of reading assessment and data analysis. There were eight items to rate the quality and utility of the course web site and five items requesting

demographic information. Four open-ended questions asked for student opinions of the best features of the course and web site and on areas in need of improvement.

Results – Phase I

The primary goal of this study was to determine educators' baseline knowledge, skills and attitudes related to using student performance data to guide decision making and the extent to which an online course in reading assessment was effective in addressing the identified gaps. Of the original 148 participants, 125 completed the study, including 57 in the treatment group and 68 in the control group. The attrition rates were 23% for the treatment group and 8% for the control group. Although a higher percentage of dropouts were teachers in the treatment group (65%) compared to the control group (50%), the mean pretest scores were comparable for both groups: 51.2 for treatment group dropouts ($n=17$) and 53.0 for control group dropouts ($n=6$).

Knowledge and Skill Development

Study question 1. Do principals/assistant principals, reading coaches, and teachers in the treatment group achieve higher pre-posttest learning gains than their counterparts in the control group?

Knowledge and skill development was measured by the TASK, a 99-item instrument designed to measure the course goals and objectives. The means and standard deviations of the pretest, posttest, and gain scores for the treatment and control groups and their respective sub-groups are shown in Table 3.

TASK pretest. The mean pretest score for the control group (55.24) was slightly higher than the treatment group's (52.62) but this difference was not statistically significant ($t = -1.70$, $df = 123$, $p > .05$), indicating similar levels of knowledge of reading assessment at the start of the

TABLE 3

Knowledge and skill development: Means and standard deviations of the TASK pretest, posttest, and gain scores

Group	<i>n</i>	Mean Score and (Standard Deviation)					
		Pretest		Posttest		Gain	
All Participants							
Treatment	57	52.62	(9.29)	76.21	(9.85)	23.59	(9.63)
Control	68	55.24	(7.97)	55.42	(7.01)	00.17	(5.52)
Principals							
Treatment	10	52.93	(8.12)	76.45	(11.07)	23.52	(8.45)
Control	12	58.03	(8.69)	57.34	(6.24)	- 0.69	(5.72)
Reading Coaches							
Treatment	11	55.85	(8.83)	79.89	(6.24)	24.04	(7.35)
Control	11	60.09	(7.36)	58.67	(6.01)	-1.42	(8.78)
Teachers							
Treatment	36	51.55	(9.72)	75.02	(10.34)	23.47	(10.70)
Control	45	53.32	(7.33)	54.11	(7.17)	00.79	(4.42)

study. Similar results were found for each sub-group (principals/assistant principals, reading coaches and teachers). Differences between treatment and control sub-group means were not statistically significant. Generally, scores were low for all three sub-groups regardless of experimental condition: 60 or less out of a total possible score of 99.

TASK posttest and gains. On the post-test, the treatment group mean was significantly higher than the control group mean ($t = 13.75$, $df = 123$, $p < .01$). Two-tailed t -tests indicate

similar results for each sub-group (principals/assistant principals, reading coaches, and teachers). For each sub-group, the treatment group mean was significantly higher than the control group mean at the .01 level.

The mean pre-post gain was higher for the participants who completed the course (treatment group) than those who did not take the course (control group) and this difference was statistically significant ($t = 16.99, df = 123, p < .01$). Again, this result was consistent for sub-groups. For each sub-group, the mean gain score was significantly higher ($p < .01$) for the treatment group than for the control group: principals/assistant principals ($t = 8.01, df = 20, p < .01$), reading coaches ($t = 7.09, df = 18, p < .01$), and teachers ($t = 13.02, df = 81, p < .01$).

Overall, the results indicate that the online course was effective in increasing the skills and knowledge required to use reading assessment data to inform instructional and curricular decisions. Moreover, the course was effective for all three sub-groups: principals/assistant principals, reading coaches, and teachers.

Reading Assessment Competencies

Study question 2: Do principals/assistant principals, reading coaches, and teachers in the treatment group show greater improvement in self-assessed competence in reading assessment than their counterparts in the control group?

The treatment ($n = 55$) and control ($n = 66$) group participants self-rated their level of competence using the Survey of Reading Endorsement Competencies (SREC). Table 4 presents the pre-treatment, post-treatment, and gain score means and standard deviations for the treatment group, control group, and each sub-group.

SREC pretest. The difference between treatment and control group means on the pre-treatment survey was not statistically significant ($t = .62, df = 119, p > .05$), indicating

equivalence of the two groups at the start of the study. Pre-post changes in self-assessed competencies were assessed by comparing the percent of participants who rated their ability as above average (i.e., above average or expert) in the 18 reading competency areas.

TABLE 4

Self-assessed competencies on reading assessment: Means and standard deviations of the SREC pretest, posttest, and gain scores

Group	n	Mean Score and (Standard Deviation)					
		Pretest		Posttest		Gain	
All Participants							
Treatment	55	54.07	(10.15)	63.22	(7.42)	9.15	(10.86)
Control	66	52.76	(12.77)	53.67	(11.62)	.91	(10.90)
Principals							
Treatment	10	50.00	(12.18)	63.60	(7.23)	13.60	(9.78)
Control	11	51.45	(9.47)	51.82	(9.24)	0.36	(6.62)
Reading Coaches							
Treatment	10	55.50	(10.52)	64.20	(9.37)	8.70	(12.76)
Control	11	54.82	(12.39)	58.00	(13.44)	3.18	(10.76)
Teachers							
Treatment	35	54.83	(9.44)	62.83	(7.07)	8.00	(10.57)
Control	44	52.57	(13.72)	53.05	(11.67)	0.48	(11.87)

Note. Treatment = online course.

Generally, the pre-treatment self-assessed competencies were low for both the treatment and control groups. In 17 of these areas fewer than 50% of participants, regardless of experimental condition, rated themselves as above average or expert (see Table 5), indicating that most did not perceive themselves to have a high level of skill at the start of the study. Both groups were most favorable about their competency in the area of Informal Assessment, although only 58% of treatment group participants and 50% of control group participants rated themselves above average in this area. Among the areas of lowest self-assessed competency (fewer than 20% of respondents) were Limited English Proficiency (LEP1 and LEP2), Students with Disabilities (SWD), Quantitative, Qualitative, Reliability and Validity.

SREC posttest and gains. The treatment group showed a higher post survey mean than the control group, which was statistically significant at the .01 level ($t = 5.27$, $df = 119$; see Table 4). The difference between the mean SREC gain scores of the treatment group ($\bar{x} = 9.15$) and the control group ($\bar{x} = .91$) was also statistically significant at the .01 level ($t = 4.15$, $df = 119$).

After course completion, the treatment group showed improvement in all 18 of the competency areas, as measured by the increase in the percent of participants who rated themselves above average. For the control group, 5 out of 18 (28%) areas showed improvement, but the increments were less than 10% of participants. The treatment group showed the greatest gain in Use and Analysis of Student Data for Progress Monitoring (36%), Informing Instruction (35%), and Differentiating Instruction (33%) in Reading.

Sub-group analysis for treatment group: The treatment group principals and teachers showed growth in their survey ratings in all 18 competency areas; the reading coaches in 16 of 18 areas (all except Validity and Limited English Proficiency 2 (LEP2)). Progress Monitoring was among the top 5 for all three groups, increasing 50% for principals, 40% for reading coaches

TABLE 5

Self-assessed competencies on reading assessment: Number, percent of ratings above average, and pre-post differences on the SREC

Competency	Treatment (n = 55)					Control (n = 66)				
	Pre	Post	Pre	Post	Diff ⁷	Pre	Post	Pre	Post	Diff
Reading Growth	19	30	35%	55%	20%	23	24	35%	36%	02%
Learner Needs	27	40	49%	73%	24%	28	28	42%	42%	00%
Formal Assessment	25	39	46%	71%	25%	29	26	44%	39%	- 05%
Informal Assessment	32	40	58%	73%	15%	33	33	50%	50%	00%
CRT	13	30	24%	55%	31%	19	17	29%	26%	- 03%
NRT	16	33	29%	60%	31%	22	16	33%	24%	- 09%
Validity	11	21	20%	38%	18%	16	11	24%	17%	- 08%
Reliability	9	16	16%	29%	13%	12	10	18%	15%	- 03%
Derived Scores	14	22	26%	40%	15%	20	16	30%	24%	- 06%
Progress Monitoring	15	35	27%	64%	36%	20	15	30%	23%	- 08%
Differentiation	18	36	33%	65%	33%	22	27	33%	41%	08%
Portfolio	19	33	35%	60%	26%	13	19	20%	29%	09%
LEP	4	13	07%	24%	16%	7	5	11%	08%	- 03%
LEP2	6	13	11%	24%	13%	7	5	11%	08%	- 03%
SWD	11	23	20%	42%	22%	10	14	15%	21%	06%
Quantitative	9	25	16%	45%	29%	10	10	15%	15%	00%
Qualitative	6	22	11%	40%	29%	11	12	17%	18%	02%
Inform Instruction	17	36	31%	66%	35%	23	22	35%	33%	- 02%

⁷ Diff = Difference Score

and 31% for teachers. All of the sub-groups showed minimal increases (20% or less) in Informal Assessment, Reliability and LEP.

Generally, principals showed the greatest number and magnitude of increases in the percent with above average ratings, although the relatively small sample size should be noted. From pre to post, this percentage increased by 30% or more on 14 of the 18 areas – all except Formal Assessment, Informal Assessment, Reliability and LEP (see Table B1 in Appendix B)

After principals, reading coaches ranked second in the number and magnitude of increases in the percent with above average ratings (see Table B2 in Appendix B), although the relatively small sample size should be noted for this sub-group also. This percentage increased by 30% or more in 8 of 18 areas, with greatest improvement in Informing Instruction (60%) and lesser but substantial gains (30% to 40%) in seven additional areas: Progress Monitoring, CRT , Formal Assessment, Learner Needs, Derived Scores, Students with Disabilities and Portfolio. In the LEP2 area, this sub-group actually declined from the pre to post measurement.

Teachers showed substantial improvement (30% or more) in only 4 of the 18 categories, as measured by pre-post increases in the percent with above average self-ratings. Those areas included Progress Monitoring, Differentiation of Instruction, Quantitative and Qualitative which showed increases between 30% and 34% (See Table B3 in Appendix B).

The SREC results indicate that participants who completed the course (treatment group) on average perceived themselves as having a higher level of competency in reading assessment than participants who did not take the course (control group). The improvements in self-assessed competencies were greatest for principals, followed by reading coaches and then teachers, as measured by increases of 30% or more in each of the 18 areas. All three groups had strong improvements in Progress Monitoring and minimal improvement (20% or less) in Informal

Assessment, Reliability and LEP. Beyond these commonalities, sub-groups varied in the areas and degree of improvement.

Stages of Concern

Study question 3. Do principals/assistant principals, reading coaches, and teachers in the treatment group show different changes in their level of concerns about reading assessment than their counterparts in the control group?

SCRRAT. The treatment ($n = 55$) and control ($n = 61$) group participants rated their level of concern about reading assessment by completing Survey of Concerns Related to Reading Assessment Training (SCRRAT). Average raw scores were calculated and converted to percentiles using a chart provided in the *Measuring Stages of Concern about the Innovation: A Manual for Use of the SoC Questionnaire* (Hall, George & Rutherford, 1979/1998). Figures 3 – 6 show the pre/post survey results illustrating the profile patterns of each group's stages of concern about reading assessment.

The seven stages of concern are *Awareness, Informational, Personal, Management, Consequence, Collaboration, and Refocus* (see Appendix C). Before the course, *Information* was the highest stage score for both the treatment (see Figure 3) and control (see Figure 4) groups with all of the participant groups scoring at or above the 90th percentile. These scores indicate that the participants had a general awareness of reading assessment and were concerned about learning more about its characteristics and use. Treatment group principals and reading coaches had two peaks, in both *Information* and *Collaboration* indicating they also had concerns about coordinating the use of reading assessments with colleagues.

Figure 3

Graph of Pre-course Stages of Concern for the Treatment Group

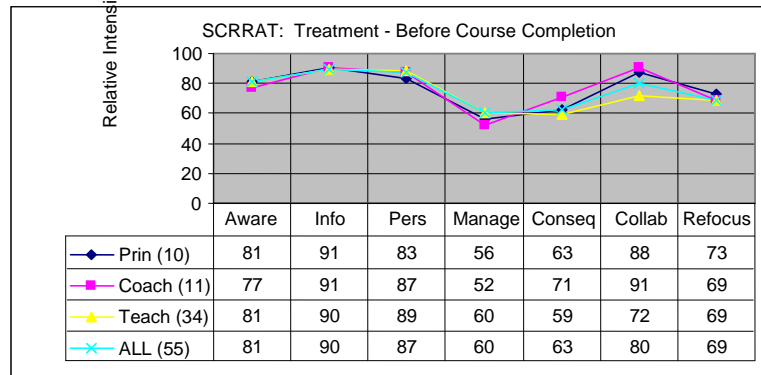


Figure 4

Graph of Pre-course Stages of Concern for the Control Group

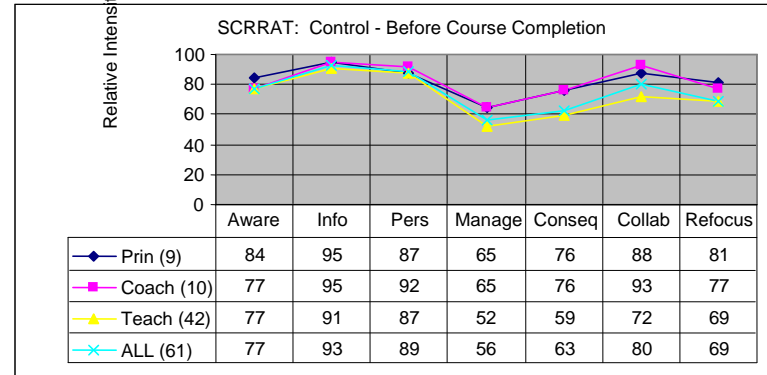


Figure 5

Graph of Post-course Stages of Concern for the Treatment Group

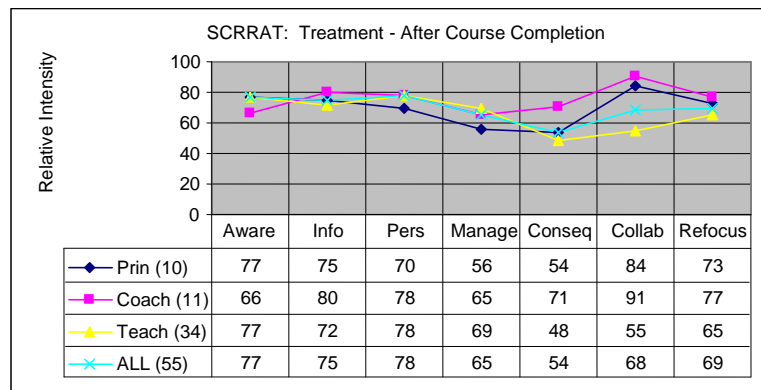
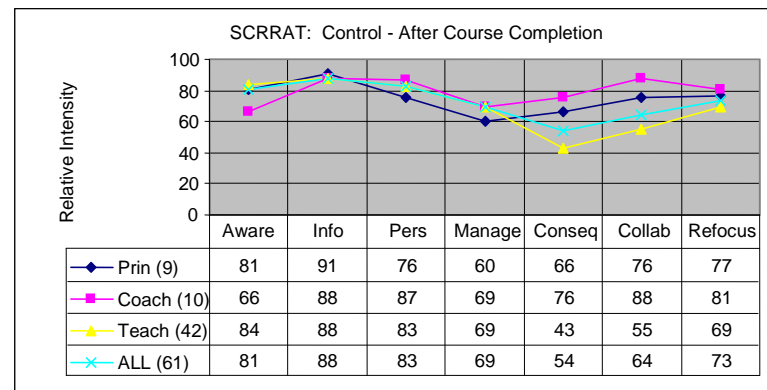


Figure 6

Graph of Post-course Stages of Concern for the Control Group



After approximately four months (post-course), the same participants completed the SCRRAT again to measure any changes in their levels of concern. The peaks of the treatment group profiles changed as illustrated in Figure 5. For the teacher group, *Personal* emerges as the major Stage (78th percentile). The principals changed to *Collaboration* at the 84th percentile as their highest Stage. The Stage of *Collaboration* continues to be reading coaches' main area of concern (91st percentile). For the control group principals and teachers, *Information* remains the highest peak at or above the 88th percentile, indicating no major change in their level of concern about reading assessment. The two equal peaks for the control group reading coaches remained at the stages of *Collaboration* and *Information* at the 88th percentile, indicating no major changes in their level of concern about reading assessment (see Figure 6).

The SCRRAT results indicate that principals/assistant principals, reading coaches, and teachers who took the online course show more change in concerns about reading assessment than those who did not take the course. While the control group remained relatively static, treatment group participants advanced to developmentally higher levels of concern.

Course Effectiveness

Study question 4. Do course participants perceive the online course on reading assessment effective in teaching the knowledge and skills required to use data for decision making?

One method of assessing the effectiveness of a course in teaching requisite knowledge and skill is by a participant evaluation of course quality. The participants in the treatment group evaluated the course quality with the 36-item Course Evaluation Survey (CES) by rating positively worded statements about the course on a five-point Likert-type scale ranging from 1 (*strongly agree*) through 5 (*strongly disagree*). To aid in the interpretation, the number of responses from the

(*strongly agree*) and the (*agree*) categories were added to create one variable for the analysis. The (*disagree*) and the (*strongly disagree*) categories also were combined.

Fifty-seven students completed the online course and 37 (65%) of them elected to complete the CES. Overall, the participants evaluated the course as effective in design, content and skill development in reading assessment (see Appendix D). As shown in Table D1, the students agreed that the course design had clear directions for tests and surveys (78%), and that they were able to understand the course goals and objectives (76%). They also agreed that the course helped them learn about the concepts of applied data analysis (76%) and assessment (70%).

In regard to future application of reading assessment (see Table D2 in Appendix D), a substantial majority of the students reported that the skills they learned in the course would help them to apply reading components in their class/school (70%), analyze and interpret data (73%), and apply data to make instructional decisions about reading (78%). Most students (78%) also reported that the course knowledge would help improve communication about pupil progress with parents and colleagues.

Weaknesses, however, were noted in the clarity and organization of course content (see Table D1). Fewer than half of respondents agreed that the content was clear and accurate (43%) and well-organized (32%) and that the directions for applications (assignments) were clear (41%). Responses to technical components of the course (see Table D3 in Appendix D) indicated low levels of agreement with ease of navigation (30%) and submittal of course assignments (41%), although 70% of the students agreed that they received satisfactory help if they had technical problems. In the open-ended section of the survey, students reported difficulty navigating from one site to another because there were “too many sites.”

The majority of students did not rate their online learning experience positively. Only 41% agreed that it met their expectations and only 35% agreed that they would like to take more online courses. Few agreed that they would recommend this course (43%, Table D2) or this online course (38%, Table D3) to a colleague or friend. Taking the above findings into account, it appears that students are responding primarily to problems related to instructional design rather course content or the course's value to their work in the school or classroom.

Discussion – Phase I

Amid growing pressure on school administrators, teachers, and reading coaches to demonstrate the effectiveness of their instructional programs and methods in terms of improved student performance, little attention has been given to whether they possess the requisite skills to do so. The present investigation documented the very limited facility of these key players with understanding and interpreting the results of reading assessments, tools essential to data-driven instructional decision making. However dismaying such deficiencies might be, this study demonstrates evidence of them and thereby provides a crucial first step toward their remediation. The second major contribution of this research was its demonstration that a readily accessible online course targeting related skills could be developed for, disseminated to, and utilized by teachers, principals and reading coaches. Finally, it was shown that such a course could achieve improvements in participants' knowledge and self-assessed competence.

The results of the pretest and pre-survey provide evidence that the training educators receive in assessment literacy is inadequate, as others have suggested (Airasian & Jones, 1993; Shafer & Lissitz, 1987; Stiggins & Conklin, 1988). The participants' (a) low pretest scores, (b) below average self-ratings on most of the reading competency items, and (c) early stages of concern about reading assessment training, provide corroborating evidence and help identify

specific areas of need for reading assessment knowledge and skill training. While the pretest results indicated that principals, teachers and reading coaches in both the treatment and control groups lacked essential knowledge, skills and attitudes to use data effectively, significant differences were found on the three measures following the successful completion of the online applied data analysis course.

After completing the course, the treatment group demonstrated significantly higher levels of reading assessment knowledge and reading assessment competency skills than their control group counterparts. The treatment group results indicated a shift to higher levels of concern about reading competency training. The results of the pre- posttest and surveys establish the effectiveness of the online course in increasing reading assessment knowledge, self-assessed levels of competence in reading assessment skills and changing the level of concerns about reading assessment. The results were similar overall and for each of the principal, reading coach and teacher sub-groups. The positive results of the course evaluation provide further support of course effectiveness. These findings have important implications for university faculty who develop courses designed to prepare new principals, reading coaches and teachers and for those who design and offer professional development for practitioners. These positive findings also contribute to the existing body of research on online teaching and learning, which has been lacking in experimental studies and triangulation of measures (Tallent-Runnels et al., 2006).

Further investigation is required to address the most salient questions associated with this topic. To what extent does enhanced knowledge, skills and attitudes related to data-driven decision making transfer to practice? Is there a difference between the treatment and control group participants' actual use of data from reading assessments for instructional decision making? What factors support or hinder the implementation and use of reading assessments?

Most importantly, if practice is transformed by possessing the requisite knowledge and skills, what are the effects of using reading assessment to improve instructional decision making on student performance? In order to answer questions related to knowledge transfer and its impact on student performance in reading, follow-up investigations are needed to assess the extent to which practice has changed. Further, an analysis of gains in performance on FCAT Reading for students associated with both treatment and control participants is required to evaluate the utility of data-driven decision making and its effects on student performance.

Limitations of the ADAPT Phase I should be considered when interpreting the results. For example, while the TASK instrument has acceptable evidence of reliability and content validity, it has not undergone the rigors of an instrument validation study. The same is true of the SREC instrument. In both cases, further study of the instruments, especially with larger and more diverse populations from different school levels and geographic locations are needed. The interpretations of the course evaluation survey results are limited by the small sample size of 37 students completing the survey garnering a return rate of 65%. Further, in that the only content area addressed was reading, we do not know the extent to which these findings might be generalized to mathematics, science, social studies or other curricular areas.

PHASE II Follow-up Study

In Phase II of the Applied Data Analysis for Principals and Teachers (ADAPT) project, a follow-up survey was conducted to determine the extent to which the participants in the Applied Data Analysis for Educational Leaders (ADAEL)⁸ treatment and control groups employed data analysis techniques in their schools or classrooms to inform data-driven instructional and curricular decisions in reading.

Research Objectives and Questions

Specifically, the ADAPT Phase II study sought to answer the following questions following the completion of the course by treatment group participants:

1. To what extent do teachers/reading coaches/principals use data to inform instructional and curricular decisions after t?
2. Is there a difference between the uses of data between participants who completed the course (treatment) and those who did not receive that training?
3. What helps/impedes the use of data for instructional decision making in schools and classrooms?

Methodology – Phase II

Participants

The 125 middle school principals, reading coaches and teachers who completed Phase I of the ADAPT⁹ project were invited to take part in the follow-up study. Participation in the study was voluntary and by informed consent. Each participant who completed the survey received an incentive payment of \$30. The response rate was 54%. The 67 respondents included 31 from the

⁸ ADAEL was name used for Phase I of ADAPT and it was the name of the on-line course (treatment).

⁹ Phase I results are presented earlier in this paper.

treatment group ($n = 31$) and 36 from the control group ($n = 36$)¹⁰ Also identified by their position, participants included 12 middle school principals/assistant principals, 13 reading coaches, and 42 teachers (Table 6).

TABLE 6
Number of Phase II Study Participants by Group and Position

Position	Number in Group		Total
	Treatment	Control	
Principals/Assistant Principals	7	5	12
Reading Coaches	7	6	13
Teachers	17	25	42
TOTAL	31	36	67

Research Instrument

The instrument administered was the Use of Data Analysis in the Classroom (UDAC) survey, developed specifically for this project to measure the extent to which principals, reading coaches and teachers used data analysis and planned to use data analysis in the future. There were six forms of the web-based UDAC survey, three survey forms for each section of the treatment group (principals, reading coaches, and teachers), and three forms for each section of the control group.¹¹ Survey questions were developed based on the Phase II research questions and results of the Phase I course evaluation survey. The survey contained multiple-choice and

¹⁰ The original 25 schools were randomly assigned by matched pairs based on school characteristics of (a) eligibility for free/reduced price lunch and (b) 8th graders scores on 2004 FCAT-SSS in reading, to either the treatment ($n = 12$ schools) or the control ($n=13$ schools) group. From the 25 schools there were 125 individual participants including 57 treatment and 68 control group members.

¹¹ Copies of the instruments are available from the author by request.

yes/no response questions divided into two major categories: participants' current/immediate and past use (items 1 – 6) and future use of data analysis components and activities (items 10 – 16). Items 7-9 inquired about participants' level of confidence in their ability to use reading plans, FCAT data for screening, and progress monitoring. Participant beliefs about the importance of principals, reading coaches and teachers in supporting the use of data analysis and application of data-driven decision making were the subject of items 17-18. The open-ended questions elicited information about the factors that participants believed would facilitate or inhibit the use of data analysis to make informed instructional decisions.

The control group surveys were similar to the treatment group surveys with the exception of two additional questions about prior coursework designed to satisfy requirements for Florida Reading Competency 3 and five additional demographic questions. The survey was created by faculty instructors of the Applied Data Analysis for Educational Leaders (ADAEL) online course that was offered as the treatment in Phase I the ADAPT project. To control content validity, expert reviews were done by the ADAPT principal investigator (also the ADAEL instructor of record), the course mentors, and the ADAPT project manager.

Procedure

The survey was posted on the Florida State University Blackboard 6 (Bb6) e-Learning System (Blackboard Inc., 2004) site that was used for the course survey and test administrations in Phase I of the ADAPT study. The survey was available for participants May 5 – May 22, 2005.

Participants received letters sent to their schools by postal mail and to their email addresses. The correspondence contained the participant user name and password to access Bb6. Once in the site, the participants were directed to choose the appropriate survey for their group

(treatment or control) and their position (principal, reading coach, or teacher). The survey required about 20-30 minutes to complete. The participants' answers to the survey questions were confidential, identified only by a subject code number.

On May 16, participants who had not submitted a survey received an email reminder. When the additional demographic questions were added to the surveys, emails with those questions were sent to the participants who had already submitted the original survey.

Results – Phase II

The participants' answers to the survey questions are confidential, identified only by a subject code number and reported as group results, not individual responses. Data analysis included the use of procedures for descriptive statistics using Microsoft Excel (2003) software and an SPSS (2003) program. The UDAC survey results for treatment and control group are shown in Table 7. Results by sub-group (principals/assistant principals, reading coaches, and teachers) are in Appendix E.

The survey contained four questions on participants' self-reported use of data and reading assessment based on a rating scale of (a) often, (b) sometimes, and (c) rarely. More than 90% of the treatment group participants who responded to this section of the survey ($n = 31$) indicated that they often or sometimes used these data and reading assessment tools. Specifically, only one (3%) treatment group member (reading coach) reported rarely using reading assessment plans and only two (6%) reported that they rarely used reading assessment data to make instructional changes. Three of the treatment group participants (9%) rarely used reading probes whereas 17 (47%) members of the control group reported that they rarely used reading probes. While 16 (50%) of the treatment group participants reported that they often led/participated in study groups, 7 (19%) control group participants reported that they did the same.

TABLE 7

Percent and difference scores on the UDAC by treatment and control groups¹²

Question	Response Options	Treatment		Control		Difference
		<i>n</i>	%	<i>n</i>	%	
1. Helped teachers develop and use reading assessment plans (Used reading assessment plans)						
	Often	13	41%	13	36%	05%
	Sometimes	17	53%	20	56%	-02%
	Rarely	1	03%	3	08%	-05%
2. Used reading assessment data to make instructional changes (Used reading assessment data)						
	Often	15	47%	15	42%	05%
	Sometimes	14	44%	17	47%	-03%
	Rarely	2	06%	4	11%	-05%
3. Led teachers in study groups for data analysis (Participated in study groups)						
	Often	16	50%	7	19%	31%
	Sometimes	13	41%	17	47%	-07%
	Rarely	2	06%	9	25%	-19%
4. Used FCAT data in 2004-05						
	Yes	30	94%	36	100%	-06%
	No	1	03%	0	00%	03%
5. Used scatter plots in 2004-05						
	Yes	14	44%	6	17%	27%
	No	17	53%	30	83%	-30%

Table continues

¹² Results may not equal 100% due to rounding.

Question	Response Options	Treatment		Control		Difference
		<i>n</i>	%	<i>n</i>	%	
6. Used reading probes						
	Often	9	28%	8	22%	06%
	Sometimes	19	59%	11	31%	29%
	Rarely	3	09%	17	47%	-38%
7. Confident in ability to lead teachers to use reading plans (Confident in ability to use reading plans)*						
	Yes	20	65%	6	55%	--
	No	0	00%	0	00%	--
	Need more training	10	32%	5	45%	--
8. Confident in ability to help teachers use FCAT for screening (Confident in ability to use FCAT for screening)						
	Yes	29	91%	28	78%	13%
	No	2	06%	1	03%	03%
	Need more training	0	00%	7	19%	-19%
9. Confident in ability to lead teachers in progress monitoring (Confident in ability to conduct program monitoring)						
	Yes	29	91%	25	69%	21%
	No	0	00%	0	00%	00%
	Need more training	2	06%	11	31%	-24%
10. Plan to help teachers use data (Plan to use data)						
	Definitely	29	91%	29	81%	10%
	Probably	2	06%	6	17%	-10%
	Unlikely	0	00%	1	03%	-03%

Table continues

Question	Response Options	Treatment		Control		Difference
		<i>n</i>	%	<i>n</i>	%	
11. Plan to help teachers use reading plans (Plan to use reading plans)						
	Definitely	29	91%	27	75%	16%
	Probably	2	06%	8	22%	-16%
	Unlikely	0	00%	1	03%	-03%
12. Plan to use 2005 FCAT data						
	Definitely	28	88%	31	86%	01%
	Probably	2	06%	3	08%	-02%
	Unlikely	1	03%	2	06%	-02%
13. Plan to use scatter plots						
	Definitely	11	34%	6	17%	18%
	Probably	14	44%	15	42%	02%
	Unlikely	6	19%	15	42%	-23%
14. Plan to help teachers use progress monitoring (Plan to set up progress monitoring)						
	Definitely	25	78%	28	78%	00%
	Probably	6	19%	6	17%	02%
	Unlikely	0	00%	2	06%	-06%
15. Plan to encourage teachers to use probes (Plan to use probes)						
	Definitely	22	69%	14	39%	30%
	Probably	8	25%	16	44%	-19%
	Unlikely	1	03%	6	17%	-14%

* 1 missing case; Control group did not have this question.

Table continues

Question	Response Options	Treatment		Control		Difference
		<i>n</i>	%	<i>n</i>	%	
16. Plan to lead teachers in study groups (Plan to participate in study groups)						
	Definitely	23	72%	21	58%	14%
	Probably	8	25%	12	33%	-08%
	Unlikely	0	00%	3	08%	-08%
17. Teacher (Principal) involvement in promoting and supporting the use of data to drive instructional decision making to improve student performance is...						
	Very Important	27	84%	27	75%	09%
	Important	4	13%	8	22%	-10%
	Not Important	0	00%	1	03%	-03%
	Irrelevant	0	00%	0	00%	00%
18. Reading coach involvement in promoting and supporting the use of data to drive instructional decision making to improve student performance is...						
	Very important	27	84%	22	61%	23%
	Important	4	13%	10	28%	-15%
	Not important	0	00%	4	11%	-11%
	Irrelevant	0	00%	0	00%	00%

Every participant except for one treatment group teacher (see Appendix E) reported that they used FCAT data in the 2004-2005 school year. Interestingly, the treatment group participants seemed divided with 44% who used scatter plots, and 53% who did not. However, 30 (83%) of the control group participants had not used scatter plots in 2004-2005. As might be expected, the control group selected the option of “need more training” more often than the treatment group. For the treatment group, 91% felt confident in both their ability to use FCAT for screening and to lead teachers in progress monitoring. None of the treatment group members

felt they needed more training in screening, and only two (6%) reported the need for more training in progress monitoring. Conversely, in the control group, 19% felt they needed more training to use FCAT scores for screening, and 31% felt that they needed more training in leading progress monitoring.

The seven questions related to plans for the use of data listed 3 response options: definitely, probably, or unlikely. In their responses to these questions, a majority ($\geq 75\%$) of both the treatment and control group members stated that they “definitely” planned to (a) use data, (b) use reading plans, (c) use FCAT data, and (d) set up progress monitoring in the future.

The areas with the largest difference scores between the treatment and control group responses are as expected given that the treatment group members received training and the control group members did not. For example, the difference score between treatment and control group responses about participating “often” in data analysis study groups was .31 and for using scatter plots in 2004-2005 was .27, indicating that the treatment group used these tools more often than their control group counterparts did. Likewise, the control group was less likely to plan to use scatter plots in the future than the treatment group members, as indicated by a difference score of -.23.

As opposed to the treatment group, the control group rarely used reading probes, with a difference score of (-.38). As one would expect, the treatment group plans to “definitely” use reading probes in the future was higher than the control group, as evidenced by a difference score of .30. The difference score for having confidence in ability to conduct progress monitoring was .21, indicating that the treatment group was more confident in their ability than the control group. Further, the difference score for “need more training” in progress monitoring

of -.24 shows that the control group felt a greater need for more instruction than the participants who had completed the course.

A qualitative analysis was conducted on the open-ended survey items by two reviewers who coded responses and established inter-rater reliability (>.90). The analysis produced eight major categories related to factors that would help or hinder the use of data analysis in the classroom (see Table E2).

The participants in the treatment group indicated that training was a factor that facilitated implementation (18%) and lack of training (15%) would be a factor that would hinder implementation. The control group also reported that training was a major attribute, with 11% listing it as needed to facilitate and 16% listing the lack of training as a hindrance. The coding definition for training included references to workshops, conferences, and in-service staff development.

Other major factors that participants noted as facilitating their implementation of data use include courses in reading competency skills, training in assessment, and instructional support. Twenty-three percent of the treatment group and 16% of the control group wrote that coursework in the Florida Reading Endorsement Competency 3 (FREC-3), especially providing resources, information, and strategies can help make the use of data analysis in the classroom possible. The need to have ready access to data was characteristic of the variables that made up the assessment category. For this category, 16% of the treatment group and 8% of the control group described as necessary. The control group comments (13%) noted that support, such as “a principal who understands and supports the use of data to inform instruction” would facilitate implementation.

The factors identified by participants that hinder their use of data for instructional decision making in schools and classrooms include the categories of access to data, personnel,

resources, and time. The lack of time was a key obstacle listed by 14% of the control group and 36% of the treatment group participants. Examples of factors listed in the time category included: (a) delayed results do not allow enough time for review, (b) lack of time to collect, analyze and/or assess data and plan accordingly, (c) lack of time to collaborate, meet, work together to discuss and use data, (etc.). Described by 12% of the control group participants, the lack of appropriate staff/personnel and adequate resources (computers, equipment, materials, etc.) each present hurdles to effective use of data for instructional decision making. The treatment group (10%) and control group (7%) also note access to data as a potential hindrance to implementation, citing problems like timely distribution/receipt of results and difficulties with accumulating data.

Discussion. The results indicate that the principals, reading coaches, and teachers who completed the course subsequently used data to inform instructional and curricular decisions. Notably, more than 90% of the treatment group participants reported that they sometimes or often used (a) reading assessment plans, (b) reading assessment data, (c) FCAT data, (d) study groups for data analysis and (e) reading probes. Approximately, 44% of the treatment group participants used scatter plots. Further, 91% felt confident in their ability to use FCAT for screening and their ability to conduct progress monitoring.

The majority of both the treatment and control group participants stated they “definitely” planned to use FCAT and other reading assessment data, reading plans, and progress monitoring in the future. The similarities in these responses indicate that educators are responding to the intensified demands to employ data-driven decision making when implementing school-based instructional programs (Institute for Educational Leadership, 2000; Johnson, 2002; Massell, 2000) despite gaps in their preparation to do so. Further, as stated earlier, federal- and state-

mandated accountability systems demand evidence for student learning outcomes (Murphy & Beck, 1994).

In some areas, there was a marked difference in data use between participants who completed the course and those who did not. Treatment group members were more likely to indicate that they participate in study groups for data analysis, use scatter plots, and plan to use reading probes. Not surprisingly, they were also more likely to express confidence in their progress monitoring skills than the members of the control group. These responses indicate the participants in the treatment group in fact applied the knowledge and skills they demonstrated on the outcome measures reported in ADAPT Phase I.

The participants described four main factors that would help with the use of data analysis in the classroom: training, courses in reading competency skills, assessment skills, and administrative support. They also indicated that lack of training, time, personnel, resources, and access to data are key hindrances to the implementation of using data for instructional decision making in schools and classrooms. While the more general factors are frequently cited as challenges in implementing educational reforms, these results identify factors that can be addressed to facilitate or block the effective use of student performance data in decision making.

Conclusions. The results of ADAPT Phase I and II make important contributions to our understanding of the challenges educators face as they attempt to use student performance data to drive decision making. They indicate online courses can be an effective device in addressing gaps in related knowledge and skills. The results of the UDAC study suggest several topics for future research. These include the need for research directed at enhancing the facilitating variables and reducing those that hinder the use of data analysis. The key focus for future

research in this area, however, will determine the extent to which understanding and using data to drive decision making results in improved student achievement in reading.

Limitations. Several factors may limit the interpretation of the results of this study. First, the course was scheduled from August through December 2004 in accordance the regular class schedule of Florida State University. The participating schools in the study were at multiple locations around Florida. During the online ADAEL course, Florida experienced one tropical storm (Bonnie) and three major hurricanes (Charley, Ivan, and Jeanne) during August and September of 2004. According to the National Weather Service (2004) (<http://www.nhc.noaa.gov>), these storms caused extensive damage to many areas of the state. All of the 12 treatment and 13 control group schools were affected to some degree by the storms including factors such as location within a storm watch or warning area and resulting storm damage (power outages, school closings, destruction of property, etc.).

In order to accommodate the needs and requests of the study participants involved in these major storm areas, the instructors issued new schedules with extended due dates three separate times and on an individual basis. The adjustments to the course schedule for treatment group students were aided by the flexible nature of the Bb-6 system. While it may be difficult to assess the total effect (if any) of these storms on the study results, we used the variable of the number of days a school was closed as an indicator of storm effect. The number of days closed for the participating schools ranged from 2 to 28. However, there was not a significant difference ($t = -.23, >.05$) between the average number of days the treatment (mean = 6.08) and the control (mean = 6.63) group schools were closed due to the storms.

References

- Ackley, D. (2001). Data analysis demystified. *Leadership*, 31(2), 28-38.
- Airasian, P. W., & Jones, A. M. (1993). The teacher as applied measurer: Realities of classroom measurement and assessment. *Applied Measurement in Education*, 6(3), 241-254.
- Anderson, N. & Henderson, M. (2004). e-PD: blended models of sustaining teacher professional development in digital literacies. *E-Learning*, 1(3), 383-394.
- Bellamy, G. T., Crawford, L., Marshall, L. H. & Coulter, G. A. (2005). The fail-safe schools challenge: Leadership possibilities from high-reliability organizations. *Educational Administration Quarterly*, 41(3), 383-412.
- Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., Walseth, P. A., Fiset, M., & Huang, B. (2004). How does distance education compare to classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research*, 74(3), 379 – 439.
- Bernhardt, V. L. (2003a). No schools left behind. *Educational Leadership*, 60(5) 26-30.
- Bernhardt, V. L. (2003b). *Using data to improve student learning in elementary schools*. New York: Eye on Education.
- Bracey, G. W. (2000). *Thinking about tests and testing: A short primer in “assessment literacy.”* Washington, DC: American Youth Policy Forum
- Blasé, J., & Blasé, Jo. (2000). Effective instructional leadership: Teachers’ perspectives on how principals promote teaching and learning in schools. *Journal of Educational Administration*, 38(2), 130-141.
- Calhoun, E. F. (1994). How to use action research in the self-renewing school. Alexandria, Virginia: Association for Supervision and Curriculum Development.

- Chen, J., Salahuddin, R., Horsch, P., & Wagner, S. L. (2000). Turning standardized test scores into a tool for improving teaching and learning: An assessment-based approach. *Urban Education, 35*(3) 356-384.
- Creighton, T. B. (May-June, 2001). Data analysis and the principalship. *Principal Leadership, 1*(9), 52-57.
- DeLong, J. & Wideman, R. (1999). Moving toward more data-based decision-making in education. *Ontario Action Researcher, 2*(2). Retrieved November 14, 2004 from <http://www.nipissingu.ca/oar/vol-2-99-20/v22e.htm>.
- Drake, T. L. & Roe, W. H. (2003). *The principalship*. Upper Saddle River, NJ: Merrill Prentice Hall.
- Florida Department of Education (2004). *Reading Scores: Statewide Comparison for 2001 to 2004*. Retrieved March 16, 2006, from <http://fcat.fldoe.org/2004/default.asp>
- Florida Department of Education (2002). *Professional development system evaluation protocol: Protocol system - final*. Retrieved February 7, 2005, from <http://www.nefec.org/>.
- Florida Department of Education, (n.d.). *Reading endorsement competencies*. Retrieved February 7, 2005, from <http://info.fldoe.org/dscgi/ds.py/Get/File1004/ReadingEndorsementCompetencies.pdf>.
- Florida Education Standards Commission (2003). *Educators accomplished practices: Accomplished competencies for teachers of the twenty first century* [Electronic version]. Tallahassee, FL: Florida Department of Education. Retrieved February 7, 2005, from <http://www.firn.edu/doe/dpe/publications/accomplished4-99.pdf>.
- Fox, D. (November, 2001). Guiding instruction through assessment: What principals need to know. *Leadership, 31*(2), 14-17.

- Frederickson, E., Pickett, A., Shea, P., Pelz, W., & Swan, K. (2000). Student satisfaction and perceived learning with online courses: Principles and examples from SUNY Learning Network. *Journal of Asynchronous Learning Networks*, 4 (2), 7-41. Retrieved August 13, 2004, from http://www.sloan-c.org/publications/jaln/v4n2/v4n2_fredericksen.asp
- Fuller, E. J., & Johnson, J. F., Jr. (2001). Can state accountability systems drive improvements in school performance for children of color and children from low-income homes? *Education and Urban Society*, 33(3), 260-283.
- Hall, G., George, A., & Rutherford, W. (1979/1998). *Measuring Stages of Concern about the innovation: A manual for use of SoC Questionnaire*. Austin, TX: University of Texas, Southwest Educational Development Laboratory.
- Hallinger, P., & Kantamara, P. (2000). Educational change in Thailand: Opening a window onto leadership as a cultural process. *School Leadership & Management*, 20(2), 189-205.
- Hassler, L. B., Buck, J. A., & Torgesen, J. K. (2004, April). *Preparation for high stakes reading assessments: The scope and nature of test preparation activities*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.
- Hassler, L., & Ogawa, G. (2002, June). *Let the data speak*. Presented at the Just Read, Florida! Leadership Conference, Tampa, FL.
- Hawley, W. D., & Valli, L (1999). The essentials of effective professional development. In L. Darling-Hammond, & G. Sykes (Eds.) *Teaching as the learning profession: Handbook of policy and practice*. San Francisco: Jossey-Bass Publishers.
- Heck, R. H., & Hallinger, P. (2005). The study of educational leadership and management: Where does the field stand today? *Educational Management Administration and Leadership*, 33(2), 229-244.

- Herman, J. L., Golan, S., & Dreyfus, J. (1990). *Principals: Their use of formal and informal data*. CSE Technical Report 315. Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Hess, F. M., & Kelly, A. P. (2005) The accidental principal: What doesn't get taught at ed schools? *Education Next*, 5(3), 35-40.
- Institute for Educational Leadership (October 2000). *Leadership for student learning: Reinventing the principalship*. A report for the Task Force on the Principalship. Washington, D.C.
- Johnson, R. (2002). *Using data to close the achievement gap: How to measure equity in our schools*. Thousand Oaks, CA: Corwin Press.
- Killion, J., and Bellamy, G. T. (2000). On the job: Data analysts focus school improvement efforts. *Journal of Staff Development*, 21(1). Retrieved October 10, 2003 from <http://www.nsd.org/library/jsd/killion211.html>.
- Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning*. University of Minnesota: Center for Applied Research and Educational Improvement.
- Levine, D. U., & Lezotte, L. W. (1990). *Unusually effective schools: A review and analysis of research and practice*. Madison, WI: National Center for Effective Schools Research and Development. (ERIC Document Reproduction Service ED 330 032).
- Lezotte, L.W. (2001). Revolutionary and evolutionary: The effective schools movement. Retrieved March 16, 2006, from <http://www.effectiveschools.com/downloads/RevEv.pdf>

- Mason, S. A. (2003, April). *Learning from data: The role of professional learning communities*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Massell, D. (2000). *The district role in building capacity: Four strategies*. (CPRE Policy Brief RB32). Philadelphia, PA: Consortium for Policy Research in Education.
- McNamara, J. F. (2000). Teaching statistics in principal preparation programs: Part 2. *International Journal of Educational Reform*, 9(4), 373-384.
- Miller, C. A. (2000, April). *School reform in action*. Paper presented to the American Educational Research Association Conference, New Orleans, LA.
- Moore, D. W., Bean, T. W., Birdshaw, D., & Rycik, J. A. (1999). *Adolescent literacy: A position statement*. Newark, DE: International Reading Association.
- Murphy, J., & Beck, L. G. (1994). Reclaiming a voice in discussions of the principal's role: A leadership challenge. In J. Murphy & K. S. Louis (Eds.). *Reshaping the principalship: Insights from transformational reform efforts* (pp. 6). Thousand Oaks, CA: Corwen Press.
- No Child Left Behind Act of 2001*, Pub. L. No. 107-10 115 Stat. 1425 (2002).
- Oosterhof, A. (2003). *Developing and using classroom assessments* (3rd ed.). Columbus, OH: Merrill Prentice Hall.
- Popham, W. J. (2000). *Modern educational measurement: Practical guidelines for educational leaders*, 3rd ed. Needham Heights, MA: Allyn & Bacon.
- Pounder, D., & Crow, G. (2005). Sustaining the pipeline of school administrators. *Educational Leadership*, 62(8), 56-60.

- Rothman, R. (2000). *Bringing all students to high standards: Report on national education goals panel field hearings*. Washington, DC: National Education Goals Panel.
- Sagor, R. (1993). *How to conduct collaborative action research*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Sagor, R., & Barnett, B. G. (1994). *The TQE principal: A transformed leader*. Thousand Oaks, CA: Corwin Press.
- Sammons, P., Hillman, J., & Mortimore, P. (1995). *Key characteristics of effective schools: A review of school effectiveness research*. London, England: Office for Standards in Education (ERIC Document Reproduction Services No. ED389926).
- Schafer, W. D., & Lissitz, R. W. (1987). Measurement training for school personnel: Recommendations and reality. *Journal of Teacher Education*, 38(3), 57-63. Retrieved January 21, 2005, from ERIC database.
- Scheurich, J. J., & Skrla, L. (2003). *Leadership for equity and excellence: Creating high-achievement classrooms, schools, and districts*. Thousand Oaks, CA: Corwin Press.
- Skinner, R. (2005). State of the states. *Education Week*, 24(17), 77-80.
- Skrla, L. Scheurich, J. J., & Johnson, J. (2000). *Equity-driven achievement-focused school districts: A report on systemic school success in four Texas school districts serving diverse populations*. Retrieved July 8, 2005, from <http://www.utdana-center.org/research/reports/equitydistricts.pdf>.
- Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND Education, RAND Reading Study Group.
- Stiggins, R. J., & Conklin, N. F. (1988). *Teacher training in assessment*. Portland, OR: Northwest Regional Educational Laboratory.

- Stiggins, R. (2001). The principal's leadership role in assessment. *NASSP Bulletin*, 85(621), 13-27.
- Tallent-Rummels, M. K., Thomas, J. A., Lan, W. Y., Cooper, S., Ahern, T. C., Shaw, S. M., & Liu, X. (2006). Teaching courses online: A review of the research. *Research in Review*, 76(1), 93-135.
- Thompson, M. (1998). *Distance learners in higher education*. Retrieved January 6, 2004, from <http://www1.worldbank.org/disted/Teaching/Design/kn-02.html>.
- Thornton, B., & Perreault, G. (2002). Becoming a data-based leader: An introduction. *NASSP Bulletin*, 86(630), 86-96.
- Valencia, R. R., Valenzuela, A., Sloan, K., & Foley, D. E. (2001). Let's treat the cause, not the symptoms: Equity and accountability in Texas revisited. *Phi Delta Kappan*, 83(4), 318-21,326.
- Wallace, R. C. J. (1996). *From vision to practice: The art of educational leadership*. Thousand Oaks, California: Corwin Press.
- Waters, J. T., Marzano, R. J., & McNulty, B. A. (2003). *Balanced leadership: What 30 years of research tells us about the effect of leadership on student performance*. Aurora, CO: Mid-continent Research for Education and Learning.
- White, S. (2005). *Beyond the numbers*. Englewood, CO: Advanced Learning Press.

Appendix B

Survey of Reading Endorsement Competencies

(SREC) Tables of Ratings *Above Average*

By Participant Sub-Group

Tables B1: Number, percent of ratings *above average*, and pre to post difference of treatment and control group principals for reading competency data

Tables B2: Number, percent of ratings *above average*, and pre to post difference of treatment and control group reading coaches for reading competency data

Tables B3: Number, percent of ratings *above average*, and pre to post difference of treatment and control group teachers for reading competency data

TABLE B1

Number, percent of ratings above average, and pre to post difference of treatment and control group principals for reading competency data

Competency	Treatment (<i>n</i> = 10)					Control (<i>n</i> = 11)				
	Pre	Post	Pre	Post	Diff	Pre	Post	Pre	Post	Diff
Reading Growth	2	5	20%	50%	30%	2	2	18%	18%	00%
Learner Needs	3	8	30%	80%	50%	5	1	46%	09%	-36%
Formal Assessment	5	7	50%	70%	20%	5	5	46%	46%	00%
Informal Assessment	5	6	50%	60%	10%	5	5	46%	46%	00%
CRT	2	7	20%	70%	50%	2	2	18%	18%	00%
NRT	2	8	20%	80%	60%	2	2	18%	18%	00%
Validity	0	3	00%	30%	30%	1	2	09%	18%	09%
Reliability	0	1	00%	10%	10%	0	1	00%	09%	09%
Derived Scores	2	6	20%	60%	40%	3	3	27%	27%	00%
Progress Monitoring	2	7	20%	70%	50%	2	2	18%	18%	00%
Differentiation	1	5	10%	50%	40%	1	3	09%	27%	18%
Portfolio	1	6	10%	60%	50%	1	1	09%	09%	00%
LEP	0	2	00%	20%	20%	0	2	00%	18%	18%
LEP2	1	4	10%	40%	30%	0	1	00%	09%	09%
SWD	2	5	20%	50%	30%	0	0	00%	00%	00%
Quantitative	2	5	20%	50%	30%	1	2	09%	18%	09%
Qualitative	1	5	10%	50%	40%	1	2	09%	18%	09%
Inform Instruction	2	5	20%	50%	30%	2	2	18%	18%	00%

TABLE B2

Number, percent of ratings above average, and pre to post difference of treatment and control group reading coaches for reading competency data

Competency	Treatment (<i>n</i> =10)					Control (<i>n</i> = 11)				
	Pre	Post	Pre	Post	Diff	Pre	Post	Pre	Post	Diff
Reading Growth	5	6	50%	60%	10%	4	6	36%	55%	18%
Learner Needs	7	10	70%	100%	30%	7	7	64%	64%	00%
Formal Assessment	3	7	30%	70%	40%	6	6	54%	54%	00%
Informal Assessment	7	8	70%	80%	10%	6	5	54%	46%	-09%
CRT	3	7	30%	70%	40%	7	6	64%	54%	-09%
NRT	5	7	50%	70%	20%	7	5	64%	46%	-18%
Validity	3	3	30%	30%	00%	2	2	18%	18%	00%
Reliability	2	3	20%	30%	10%	2	2	18%	18%	00%
Derived Scores	2	5	20%	50%	30%	6	3	54%	27%	-27%
Progress Monitoring	3	7	30%	70%	40%	4	5	36%	46%	09%
Differentiation	5	7	50%	70%	20%	4	7	36%	64%	27%
Portfolio	4	7	40%	70%	30%	3	6	27%	54%	27%
LEP	0	1	00%	10%	10%	2	1	18%	09%	-09%
LEP2	2	1	20%	10%	-10%	1	1	09%	09%	00%
SWD	3	6	30%	60%	30%	1	4	09%	36%	27%
Quantitative	4	6	40%	60%	20%	3	3	27%	27%	00%
Qualitative	3	4	30%	40%	10%	3	3	27%	27%	00%
Inform Instruction	3	9	30%	90%	60%	5	6	46%	54%	09%

TABLE B3

Number, percent of ratings above average, and pre-to-post difference of treatment and control group teachers for reading competency data

Competency	Treatment (<i>n</i> = 35)					Control (<i>n</i> = 44)				
	Pre	Post	%Pre	%Pos t	Diff	Pre	Post	%Pre	%Pos st	Diff
Reading Growth	12	19	34%	54%	20%	17	16	39%	36%	-02%
Learner Needs	17	22	49%	63%	14%	16	20	36%	46%	09%
Formal Assessment	17	25	49%	71%	23%	17	15	41%	34%	-07%
Informal Assessment	20	26	57%	74%	17%	22	23	50%	52%	02%
CRT	8	16	23%	46%	23%	10	9	23%	21%	-02%
NRT	9	18	26%	51%	26%	13	9	30%	21%	-09%
Validity	8	15	23%	43%	20%	13	7	30%	16%	-14%
Reliability	7	12	20%	34%	14%	10	7	23%	16%	-07%
Derived Scores	10	11	29%	31%	03%	11	10	25%	23%	-02%
Progress Monitoring	10	21	29%	60%	31%	14	8	32%	18%	-14%
Differentiation	12	24	34%	69%	34%	17	17	39%	39%	00%
Portfolio	14	20	40%	57%	17%	9	12	21%	27%	07%
LEP	4	10	11%	29%	17%	5	2	11%	04%	-07%
LEP2	3	8	09%	23%	14%	6	3	14%	07%	-07%
SWD	6	12	17%	34%	17%	9	10	20%	23%	02%
Quantitative	3	14	09%	40%	31%	6	5	14%	11%	-02%
Qualitative	2	13	06%	37%	31%	7	7	16%	16%	00%
Inform Instruction	12	22	34%	63%	29%	16	14	36%	32%	-05%

Appendix C

Definitions of the Stages of Concern about the Innovation¹³

- 0 **AWARENESS**: Little concern about or involvement with the innovation is indicated.
- 1 **INFORMATIONAL**: A general awareness of the innovation and interest in learning more detail about it is indicated. The person seems to be unworried about herself/himself in relation to the innovation. She/he is interested in substantive aspects of the innovation in a selfless manner such as general characteristics, effects, and requirements for use.
- 2 **PERSONAL**: Individual is uncertain about the demands of the innovation, her/his inadequacy to meet those demands, and her/his role with the innovation. This includes analysis of her/his role in relation to the reward structure of the organization, decision making, and consideration of potential conflicts with existing structures or personal commitment. Financial or status implications of the program for self and colleagues may also be reflected.
- 3 **MANAGEMENT**: Attention is focused on the processes and tasks of using the innovation and the best use of information and resources. Issues related to efficiency, organizing, managing, scheduling, and time demands are utmost [sic].
- 4 **CONSEQUENCE**: Attention focuses on impact of the innovation on students in her/his immediate sphere of influence. The focus is on relevance of the innovation for students, evaluation of student outcomes, including performance and competencies, and changes needed to increase student outcomes.
- 5 **COLLABORATION**: The focus is on coordination and cooperation with others regarding use of the innovation.
- 6 **REFOCUSING**: The focus is on exploration of more universal benefits from the innovation, including the possibility of major changes or replacement with a more powerful alternative. Individual has definite ideas about alternatives to the proposed or existing form of the innovation.

¹³ Figure from Hall, G. E., George, A. A., & Rutherford, W. L. (1979/1998). *Measuring Stages of Concern about the Innovation: A Manual for Use of the SoC Questionnaire*. Austin, TX: Southwest Educational Development Laboratory (p.7).

Appendix D

Course Evaluation Survey (CES) Frequency Tables

Table D1: CES Survey Frequency Table for Course Content

Table D2: CES Survey Frequency Table for Application of Course Content

Table D3: CES Survey Frequency Table for Technical Components of the Course

TABLE D1

Course Evaluation Survey (CES): Frequency table for course content

Item	Strongly Agree + Agree	Neutral	Strongly Disagree + Disagree
Content clear and accurate	43.20	21.60	35.10
Content well organized	32.40	29.70	37.80
Understood goals and objectives	75.70	10.80	13.50
Sufficient feedback	67.50	5.40	27.00
Understood criteria to evaluate work	64.80	24.30	10.80
Directions for tests and surveys clear	78.40	18.90	2.70
Directions for applications clear	40.50	24.30	35.10
Course helped me learn concepts of assessment	70.20	10.80	18.90
Course helped me learn skill in Reading Competency 3	64.80	32.40	2.70
Course helped me learn concept of applied data analysis	75.70	10.80	13.50
Internal and external links helpful	67.50	13.50	18.90
Learned about various tool on the Internet	70.00	16.20	13.50

TABLE D2

Course Evaluation Survey (CES): Frequency table for application of course content

Item	Strongly Agree + Agree	Neutral	Strongly Disagree + Disagree
Skills learned will help me apply component of Reading Components in my class/school.	70.20	21.60	8.10
Skills learned will help me analyze and interpret data.	72.90	21.60	5.40
Skills learned will help me improve my communication about student progress with parents and colleagues.	78.40	13.50	8.10
Participation provided me with the skills and information I need to apply data to make instructional decisions about reading.	78.40	10.80	10.80
I would recommend this course to a friend or colleague.	43.20	27.00	29.70

TABLE D3

Course Evaluation Survey (CES): Frequency table for technical components of the course

Item	Strongly Agree + Agree	Neutral	Strongly Disagree + Disagree
It was easy to access the course website.	62.10	21.60	16.20
It was easy to navigate the course website.	29.70	32.40	37.80
It was easy to submit course assignments.	40.50	21.60	37.80
It was easy to communicate using Blackboard email.	54.00	18.90	27.00
If I had technical problems, I received satisfactory help.	70.20	21.60	8.10
My experience in online learning met my expectations.	40.50	21.60	37.80
From my experience with this course, I would like to take more online courses.	35.10	27.00	37.80
I would recommend taking this online course to a friend or colleague.	37.80	27.00	35.10

Appendix E

Use of Data Analysis in the Classroom (UDAC) Results

Table E1: Percent and difference score for UDAC treatment and control groups by position
Table E2: Frequency and percent of use of data analysis in the classroom comment categories

TABLE E1

Percent and difference score for UDAC treatment and control groups by position

Current Use of Data	Treatment						Control					
	Principal (7)		Coach (7)		Teacher (17)		Principal (5)		Coach (6)		Teacher (25)	
	N	%	N	%	N	%	N	%	N	%	N	%
Helped teachers develop and use reading assessment plans (Used rdg assmt plans)												
Often	5	0.71	3	0.43	5	0.29	0	0.00	2	0.33	11	0.44
Sometimes	2	0.29	3	0.43	12	0.71	4	0.80	4	0.67	12	0.48
Rarely	0	0.00	1	0.14	0	0.00	1	0.20	0	0.00	2	0.08
Used rdg assessment data to make instructional changes (Used rdg assmt data)												
Often	3	0.43	5	0.71	7	0.41	2	0.40	3	0.50	10	0.40
Sometimes	3	0.43	2	0.29	9	0.53	2	0.40	3	0.50	12	0.48
Rarely	1	0.14	0	0.00	1	0.06	1	0.20	0	0.00	3	0.12
Led teachers in study groups for data analysis (Participated in study groups)												
Often	4	0.57	4	0.57	8	0.47	0	0.00	1	0.17	6	0.24
Sometimes	3	0.43	2	0.29	8	0.47	3	0.60	4	0.67	10	0.40
Rarely	0	0.00	1	0.14	1	0.06	2	0.40	1	0.17	6	0.24
Used FCAT data in 2004-05												
Yes	7	1.00	7	1.00	16	0.94	5	1.00	6	1.00	25	1.00
No	0	0.00	0	0.00	1	0.06	0	0.00	0	0.00	0	0.00
Used scatter plots in 2004-05												
Yes	3	0.43	2	0.29	9	0.53	1	0.20	0	0.00	5	0.20
No	4	0.57	5	0.71	8	0.47	4	0.80	6	1.00	20	0.80
Used reading probes												
Often	3	0.43	3	0.43	3	0.18	1	0.20	3	0.50	4	0.16
Sometimes	4	0.57	3	0.43	12	0.71	2	0.40	2	0.33	7	0.28
Rarely	0	0.00	1	0.14	2	0.12	2	0.40	1	0.17	14	0.56
Confident in ability to lead teachers to use reading plans (Confident in ability to use rdg plans) <i>Note: 1 missing case; control group did not have this question</i>												
Yes	5	0.71	4	0.57	11	0.69	3	0.60	3	0.50	--	--
No	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	--	--
Need more training	2	0.29	3	0.43	5	0.31	2	0.40	3	0.50	--	--
Confident in ability to help teachers use FCAT for screening (Confident in ability to use FCAT for screening)												
Yes	7	1.00	7	1.00	15	0.88	3	0.60	6	1.00	19	0.76
No	0	0.00	0	0.00	2	0.12	0	0.00	0	0.00	1	0.04
Need more training	0	0.00	0	0.00	0	0.00	2	0.40	0	0.00	5	0.20
Confident in ability to lead teachers in progress monitoring (Confident in ability to conduct progress monitoring)												
Yes	7	1.00	7	1.00	15	0.88	3	0.60	6	1.00	16	0.64
No	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Need more training	0	0.00	0	0.00	2	0.12	2	0.40	0	0.00	9	0.36

TABLE E2

Frequency and percent of Use of Data Analysis in the Classroom (UDAC) comment categories

	Facilitate = Presence of Factor				Hinder = Absence of Factor			
	Control		Treatment		Control		Treatment	
Factors	No.*	%	No.	%	No.	%	No.	%
Access to data	4	07	1	02	3	07	4	10
Assessments	5	08	9	16	3	07	3	08
Course (Reading Endorsement skills)	10	16	13	23	3	07		
Class size							1	03
Computer	3	05			2	05	2	05
Data analysis			5	09				
Funds					1	02	2	05
Other	1	02	1	02				
Personnel	5	08	1	02	5	12		
Program	4	07	4	07	2	05		
Reading coach	5	08	5	09	1	02		
Resources	4	07	1	02	5	12	3	08
SIP			1	02				

	Facilitate = Presence of Factor					Hinder = Absence of Factor				
	Control		Treatment		Control		Treatment			
Factors	No.*	%	No.	%	No.	%	No.	%		
Software	3	05	2	04						
Study group	1	02	3	05						
Students					2	05	1	03		
Support	8	13	1	02	2	05	1	03		
Technology					1	02	2	05		
Time	1	02			6	14	14	36		
Training	7	11	10	18	7	16	6	15		
Total Comments	61		57		43		39			

Note. No. = number of comments

