Volume 22, No. 2—Spring 2009

MID-WESTERN EDUCATIONAL RESEARCHER

Official Publication of the Mid-Western Educational Research Association

Harris-Stowe State University
St. Louis, Missouri

Inside this Issue:

Fall Conference Preview 2009

MWERA
www.mwera.org

On the Cover–Harris-Stowe University

Harris-Stowe State University's US Bank Entryway Arch, displayed on the cover, is emblematic of the great things that occur when you enter the arches. Built in 2004, the US Bank Entryway Arch really was the gateway for the momentous achievements soon to come. The following year, Harris-Stowe officially became a university; the Anheuser-Busch School of Business Administration, located on HSSU's South Campus, opened; and the former Vashon Community Center facility was named to the National Register of Historic Places by the Federal Government.

Although its university status may be new, Harris-Stowe has been educating St. Louis for more than 150 years. This historic institution is known for its distinguished roots as a Teacher Education college and for offering the most affordable baccalaureate degree in the metropolitan St. Louis area. Through the years, HSSU expanded its curriculum to include 12 challenging degree programs in the areas of Business Administration, Urban Specializations, Arts and Sciences, and of course, Teacher Education.

Currently, 2,000 students matriculate at Harris-Stowe, but students still get the individual attention they need with a low 30:1 student-to-staff ratio. Student residents live comfortably on campus in the Rev. Dr. William G. Gillespie Residence Hall and Student Center, which houses 228 students in four-bedroom suites, equipped with a common living and dining area, kitchenette and two bathrooms.

Harris-Stowe boasts seven National Association of Intercollegiate Athletics (NAIA) athletic teams and around 45 student organizations. Harris-Stowe continues to expand, and this year, the university will open the $17-million state-of-the-art Early Childhood Development and Parenting Education Center. Stay tuned as Harris-Stowe continues its goal of becoming one of the area's most flourishing universities.

Call for Manuscripts

The ***Mid-Western Educational Researcher*** is a scholarly journal that publishes research-based articles addressing a full range of educational issues. The journal also publishes literature reviews, theoretical and methodological discussions that make an original contribution to the research literature, and feature columns. There are four issues of the journal published annually.

Manuscripts are accepted from faculty, students, and professionals working in educational or non-educational settings. Membership in the MWERA is not required in order to submit a manuscript for review. The editors encourage the submission of revised papers that have been presented at the annual meetings of the MWERA, AERA, and other professional organizations.

In addition, the editors encourage quality graduate student submissions through a devoted section, "Graduate Student Research," for papers authored by graduate students (as either sole or first author). This does not preclude manuscripts authored by graduate students from appearing in the main section of the journal. The submission, review, and publication of manuscripts in this section conform to the descriptions and standards of the journal outlined below. Manuscripts should be submitted electronically with *MWER Graduate Student Manuscript* as the subject line. It is essential to identify yourself as a graduate student to be considered for the Graduate Student Section. Verification of graduate student status will be required if the manuscript is accepted for publication in *MWER*.

All manuscripts must be submitted electronically for review to Dr. Julia Matuga, Submissions Co-Editor, at mer@bgsu.edu as an email attachment. Indicate in the subject line that this is a MWER manuscript. Manuscripts should be formatted as an MS Word document using 12 point Times New Roman font. Manuscripts should conform to the style and format described in the *Publication Manual of the American Psychological Association, 5th edition*. All manuscripts should be typed, double-spaced, with 1½ inch margins on all sides, and include page numbers. An abstract of fewer than 100 words should accompany the manuscript. The author's name, contact information, and affiliation should appear on the title page only. Submissions typically are fewer than 20 pages in length, including references, title page, and abstract.

All manuscripts will be acknowledged electronically upon receipt. Please note that authors are responsible to submit manuscripts that are free of grammatical and mechanical errors. Manuscripts will be screened initially for format and fit for the journal by the editors. Appropriate manuscripts will be submitted to blind review. The editors reserve the right to make minor modifications in order to produce a more concise and clear article. Contributors acknowledge by virtue of their submission to the journal that they will consent to have their work available internationally through the EBSCO portal, as per agreement with the MWERA.

Questions regarding the journal should be directed to the Submissions Co-Editor:

Dr. Julia M. Matuga
College of Education & Human Development
Bowling Green State University
444 Education Building
Bowling Green, OH 43403
(419) 372-7317
mer@bgsu.edu

The *Mid-Western Educational Researcher* (ISSN 1056-3997) is published quarterly by the MidWestern Educational Research Association through The Ohio State University. The summer issue serves as the annual meeting program. Non-profit postage paid at Columbus, Ohio, with permission of Cheryl Achterberg, Dean, College of Education and Human Ecology, The Ohio State University.

POSTMASTER: Send address change to: Sharon McNeely, Ph.D., P. O. Box 34421, Chicago, IL 60634-0421

THE MID-WESTERN EDUCATIONAL RESEARCH ASSOCIATION

Gift Membership

A **GIFT MEMBERSHIP**
has been given to you

_____

by _____

Your name is now included as a member of one of the most recognized, well-respected educational research groups in the United States and Canada. Your one-year membership includes a subscription to the ***MID-WESTERN EDUCATIONAL RESEARCHER***, the association's journal. Members pay reduced registration fees for the annual meeting held in October. The conference attracts many nationally-recognized leaders in educational research. Enjoy your membership!

Thank you for providing your colleague, student, or friend with a special one-year gift membership to the MID-WESTERN EDUCATIONAL RESEARCH ASSOCIATION. It is a gift of professional involvement that is sure to be appreciated throughout the year. To give your gift membership, photocopy this page, fill out the bottom portion, and mail it with your check to:

Sharon McNeely, Ph.D., P.O. Box 34421, Chicago, IL 60634-0421

	Person Receiving Gift Membership	**Person Giving Gift Membership**
Name	_____	_____
Address	_____	_____
	_____	_____
	_____	_____
Affiliation	_____	_____
Work Phone	_____	_____
Home Phone	_____	_____
E-mail	_____	_____
Fax	_____	_____
Division Preference (optional)	_____	_____

Professional Membership: $45.00
Student Membership: $25.00

Make your check payable to: MWERA

The Ohio State University
2500 Kenny Road
Columbus, Ohio 43210

MID-WESTERN EDUCATIONAL RESEARCHER

Volume 22, Number 2 · Spring 2009 ISSN 1056-3997

MWER Publication Address

Julia Matuga
Bowling Green State University
444 Education
Bowling Green, OH 43403
Phone: 419-372-7317
e-mail: mer@bgsu.edu

MWER Subscription & MWERA Membership Information

Sharon McNeely, Ph.D.
P. O. Box 34421
Chicago, IL 60634-0421
Phone: 773-442-5518
Fax: 773-736-7033
e-mail: mwera_ed@yahoo.com

Instructional Approaches in Middle-Level Reading Programs: A State-Wide Profile of Focus and Support

Francine Falk-Ross
Pace University

Abstract

A survey of assessment and instructional approaches to literacy education in middle schools was distributed to coordinators of reading programs in schools state-wide in Illinois to determine the alignment between existing schools' literacy program organization and new directions mandated by researchers for instruction for young adolescents. This exploratory study used questions and statements on time commitments, focused reading approaches, forms of assessment and evaluation of reading achievement, mentoring of struggling readers, and the types of extracurricular reading opportunities for middle-level students. Results indicated unbalanced emphases in literacy development. Implications for teacher education and school reorganization are discussed.

Due to changes in school populations and results of studies indicating the special needs of students in grades 5-8 (e.g.,Combs, 2003; Unrau, 2004), there is an important need to conduct research related the nature of literacy education for middle level students, with implications for teacher education programs that focus on middle-level teacher preparation. Specific knowledge is needed of the nature of existing school programs in which new middle-level educators will be placed for practicum experiences and eventually work so that the literacy instruction and assessment for students will be consistent with national mandates. Models of effective teacher education program elements for preparing preservice and practicing middle-level educators are needed to meet the literacy (specifically, reading) needs of young adolescent students. The purpose of this exploratory survey project was to report the results of a (Midwestern) state-wide survey of middle-level reading assessment and instruction as a barometer of recent changes in programming and to gain insight into the nature of existing programs in the local areas, partnership schools, and outlying school districts. A second purpose was to propose educational applications that derive from these results in order to specify targets for further research and to develop programmatic content material in the area of middle-level reading, especially for teacher education, consistent with national middle-level certification programs.

Theoretical Framework

Research indicates that young adolescents in grades 5-8 share characteristics that differ considerably from elementary and secondary-level students, and require educational approaches that are specific to these needs (Jackson & Davis, 2000; National Middle School Association, 2003). Following mandates to reorganize junior high program structure to fit the new middle-level models, extensive studies completed in the 1990's, especially in Illinois middle schools, have shown progress in all achievement scores (including reading) and behavioral issues of students in this age group (Felner, Jackson, Kasak, Mulhall, Brand, & Flowers, 1997). However, it has been observed by this researcher during school visitations, conversational discussions, and informal questioning with middle-level educators that these research-based elements of reading assessment and instruction in middle-level grades are not consistently used; most schools are still in the process of approaching the use of core elements (e.g., collaborative learning grouping, advisory periods, or exploratory classes) for effective learning. As a result, literacy education within these middle grade schools may not consist of the prescribed components as set forth by middle-level literacy experts (e.g., Atwell, 1998; Combs, 2003; Rycik & Irvin, 2004).

A second consideration is that researchers and teacher educators are in the process of developing a certification program for teacher licensure. In preparation for focused efforts towards middle school reform, schools have been transforming their reading programs (e.g., Doda & Thompson, 2002). It is necessary to organize information and resources for perspective middle-level educators, separate from those of elementary and high school educators, that target the unique needs of middle school students, especially in the area of teacher education for literacy development.

Methodology

Participants in the survey project were chosen using the Illinois State Board of Education's (ISBE) list of middle and junior high public schools. For the purpose of this study, the middle or junior high schools did not need to be separate entities from elementary or secondary schools as long as they were identified as being a teaching organization addressing the educational needs of students in grades 5-8. Participants were limited to those in public schools, which were required to meet the state standards for teaching/learning and national mandates for education. The target population included educators who organized and directed the reading program in those schools, such as reading coaches, reading endorsement recipients, reading specialists, and reading curriculum directors. Five hundred surveys were sent to educators identified on the lists or through telephone calls to the district as being responsible for and knowledgeable about the school district's middle-level reading program; approximately 20% were completed and returned for analysis. Those educators who did not return the completed survey in a six-week period

were sent follow-up letters as reminders and a second copy of the survey, if necessary. No incentive was provided for the educators to complete and return the survey; contributing to the general knowledge of state-wide needs was held as the motivational element in completing the survey.

The study was organized to determine the context in which the reading programs were situated within the school's larger educational program and to investigate the manner in which the elements were integrated into classrooms. The four-page survey consisted of a listing of forced choice questions in the areas of general school program organization involving reading activities (exploratory cases, advisory periods, team organization, decision-making, and parent and community programs) the nature of reading lessons (texts and literacy resources, media and materials, reading components, and educators' roles), forms of assessment and evaluation of reading achievement (formative and summative data collection), and forms of mentoring for struggling readers. Responses within categories included rating in terms of frequency of use and preferences. Items were chosen by the researcher and an outside rater for reliability in the alignment of existing programs with those described in research as aligned with the new middle school movement. An open-ended question asking for additional information or topics that the participant chose to include were analyzed using a constant comparative coding scheme (Corbin & Strauss, 2007) to identify themes in the responses of these reading teachers/specialists.

A mail survey was chosen to gather data on the middle-level reading programs, although this mode of survey has been shown to draw the lowest response rate (Hox & de Leeuw, 1994); this may not influence nonresponse bias (Groves, Fowler, Couper, Lepkowski, Singer, & Tourangeau, 2004). The surveys were sent to educators with cover letters explaining the purposes and procedures for completing the survey as well as return envelopes for their responses. Contact information provided for questions about survey completion was used by 5 educators to further qualify their responses.

Results

The statistical analysis was limited to aggregating the percentages of use of specific reading strategies, to describing statistics of the reading teachers, coaches or specialists, and to describing the types of class offerings. A close consideration of the findings point to areas of program development and reading instruction that require more attention and stronger focus.

Program organization affecting reading instruction

In the area of reading program organization, results of the survey supplied by participants' feedback indicated that time allotted for grade-level meetings averaged ~70 minutes per week, allowing teachers some time to collaborate and coordinate their lessons to meet curricular requirements and students' needs. From more qualitative feedback it was noted that some small amount of the time spent in these meetings was devoted to integrating units of instruction. This is consistent with the focus of the middle-level movement on the importance of faculty teaming. Exploratory coursework, also shown to be significant in the broad academic and socio-emotional development of young adolescents, was mostly focused on art and second language learning, with some schools still occasionally offering woodshop classes. Decision-making seemed to be shared by the principal and teachers in most districts; however, the time spent was small, hinting at a lack of reflective practice as an ongoing process. Although there were work and communication with parents and community representatives, these were organized by specific faculty and or administrators, and time was not specifically set aside each week (or month) for outreach events. This was also seen by the noticeable lack of service learning experiences and partnership activities. Advisory classes, when offered, were spent mostly on resolving interpersonal issues and less on schoolwork concerns. Specifically, teachers seemed to struggle with not having enough time to address the advisory element. The purpose of advisory periods, according to research supporting middle-level reform efforts, is to assist students in determining course options and to counsel them in academic concerns, as well as be accessible for interpersonal questions. Therefore, it would be preferable to spend more time focused on school-wide issues. These time allotments can be seen in graphic form in Figure 1, below.

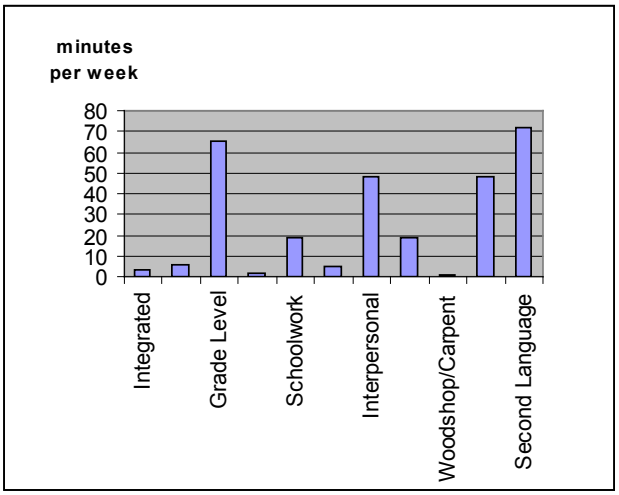

Figure 1. Reading Program Organization

Elements of the Reading Program

In the area of the nature of reading lessons, it appears from the survey that several specific areas, such as reading comprehension (45%) and vocabulary development (20%), were consistently a focus within these middle grade reading programs, while word identification (10%), fluency factors (10%), and listening skills and listening comprehension (15%) were not as well represented in reading instruction. This last factor is of concern since so much in the form of following directives for assignment completion and story or procedure sequence is expected in these grades.

Whole class lecture predominated for reading instruction (22%). Oral read alouds by students (18%) was only used a bit more than guided reading approaches (15%), while small group heterogeneous work in the form of reading/writing workshop, interdisciplinary units, and literature circles were used as frequently as independent worksheets (~10%). The use of basal anthologies or reading series (28%) occurred just slightly more frequently than use of tradebooks (26%) as a preference, which indicates a balance in use of textual resources for reading. Literacy centers, meant to continue the work completed in guided reading groups and to encourage independent thinking and innovative practice with reading, were not frequently used (6%).

The integration of media and computer programs (6-8%) as sources for learning content material and developing visual literacy were still not common in the surveyed schools. Instead, media and computers were used for independent practice literacy activities. Technology resources were utilized occasionally to complement and implement reading activities. Visual displays, such as charts, word walls, and graphic organizers, were used (8%) more than iPods and musical resources (~2%).

Input from educators to make decisions about the reading elements and approaches that were used were made mostly by the language arts/reading teachers (41%), in collaboration with other classroom teachers (15%), the reading specialist (8%), and the special education educator (10%). The educators' backgrounds are important to consider in the type of decision-making that occurs. Some of the teachers had middle-grade endorsement, fewer had reading certification or endorsements (3%), mixed numbers came with elementary and secondary experience, and many teachers received professional development programs or attended inservice presentations (11%) to develop new knowledge in the area of literacy instruction.

Formative and summative data

In the area of evaluation of students' achievement in classes, informal teacher-made tests (~31%) and questions taken from basals or reading series anthologies (~ 18%) were the most frequently used assessment tools. Portfolio collections of process learning and reading response papers, which are particularly helpful to the teacher in determining individual needs and strengths, were used approximately 10% of the time for weekly assessment activities. Running records which are helpful to teachers in identifying students' word identification knowledge, reading comprehension, and measures of fluency, were used less than 10% of the time. Running records are an important element in flexible grouping programs for guided reading instruction, and yet, from comments in the open response portion of the survey, teachers felt that these were usually used in the lower grades for younger students. Timed worksheets were used to focus and measure vocabulary knowledge and reading comprehension by teachers on an average of ~2% of the allotted assessment time per week.

In the area of mentoring struggling readers, instructional accommodations within the classroom were used most frequently (~24%), guided by collegial and, occasionally, special educators' advice. The decisions that teachers make in organizing these accommodations were usually provided through oral discussions, not in written form, and were reliant on teachers' own judgment in their classrooms. Reading support outside of the classroom (e.g., separated instruction such as special classes for ELLs and students with LD challenges, or resource rooms for individual tutoring) were used almost as frequently (~20%). Giving individual attention to struggling readers in the classroom (e.g., aides or paraprofessional assistants) was more difficult to accomplish for teachers, and this occurred with 10% frequency, equal to that amount available during guided reading opportunities within competence-based small groups.

Extracurricular reading opportunities for middle-level students included school-home content area reading and research projects (8%) or after-school activities (~10%), such as book clubs and debate team meetings. At this grade level, students are usually left to decide for themselves how they can expand their reading opportunities, and the results of this surveyed supported that practice. Community projects and service learning activities were used infrequently by these middle-level programs for reading purposes.

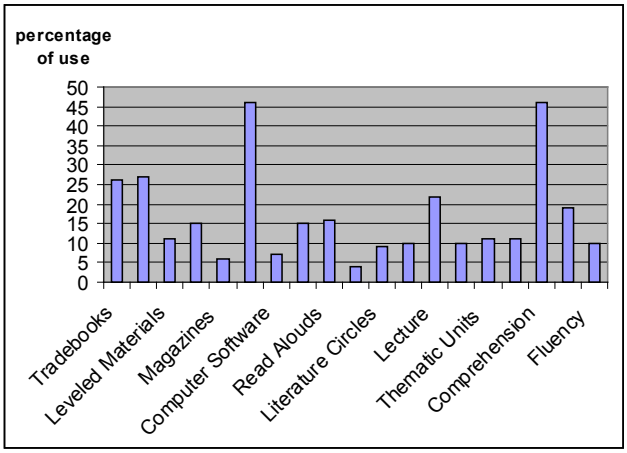

Figure 2. Elements of Reading Program

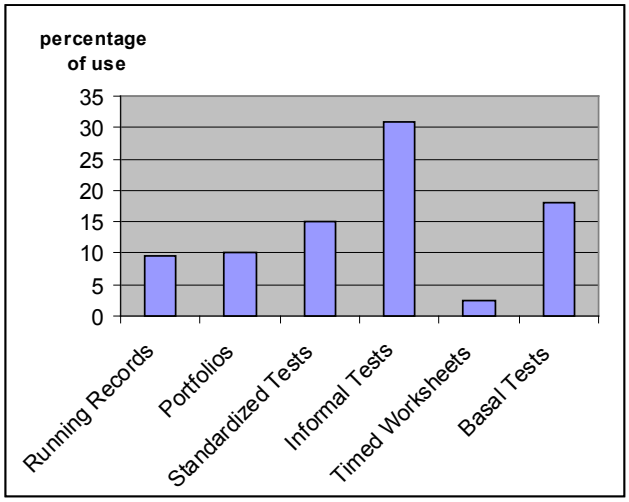

Figure 3. Formative and Summative Data Collection

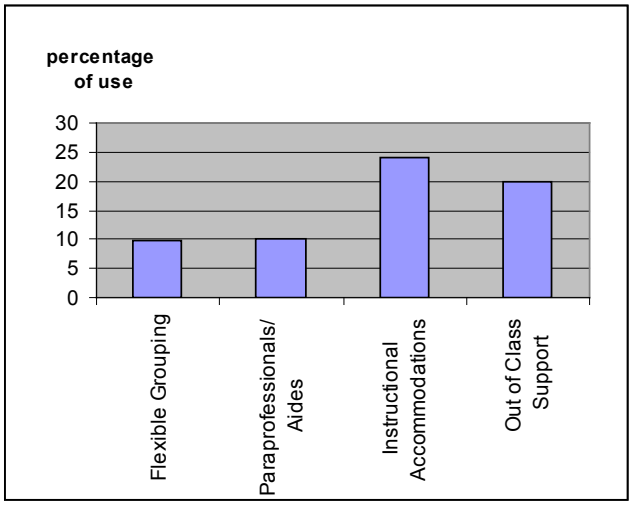

Figure 4. Mentoring of Struggling Readers

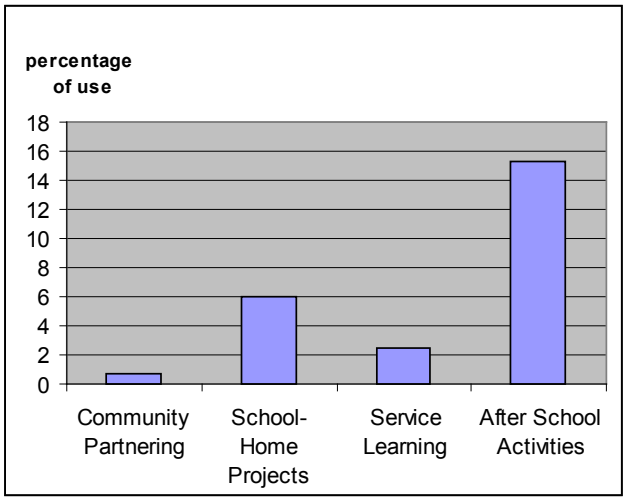

Figure 5. Types of Extracurricular Reading Opportunities

Discussion and Implications

The educational significance of this study is it allows a view into a large number of middle grade programs and provides insight into how everyday programs can be expanded to become consistent with middle school movement initiatives. It underscores the necessity for middle school educational specialists and administrative leaders to partner in reviewing programmatic elements and priorities to align with students' best interests and literacy needs within educational curricula.

In light of results from this survey, one area for reconsideration will include an increased emphasis on word identification strategies for its influence on vocabulary development in young adolescents' reading success, and on fluency strategies for comprehension. Unrau (2004) and Combs (2003) discuss the need for a focus on the approximately 25% of middle-level students who read several grades levels below their classmates and who struggle to read in content area classrooms which has been indicated in national testing results (e.g., NAEP, 1999), citing decoding skills and sight words as part of the problem. A growing number of second language learners in middle-level classrooms partially accounts for the need to more equally balance the types of reading instruction in middle grades. For example, Illinois has been identified as one of the six states that now accounts for a majority of immigrant children in the United States (Ruiz-de-Velasco & Fix, 2000).

The predominance of whole class teaching and infrequent use of guided reading and literacy center use needs to be considered in specific reading programs, especially those with a significant number of struggling readers, including second language learners. The growing diversity in classroom populations mandate the use of differentiated approaches to teaching and increased opportunities for individual conferencing with reading teachers. Flexible grouping to support the different levels and competencies of young adolescents as they learn in uneven spurts has been shown to be an effective approach to literacy learning (Allen, 2000; Fountas & Pinnell, 2001). Independent learning within structured formats such as reading/writing workshop (Atwell, 1998) or literature circles (Raphael, Kehus, & Demphousser, 2001; Redman, 1995) provide students with opportunities to socialize and share new information with peers.

The still low level of media and technology use as integrated elements in literacy education may also be of concern to middle-level educators. Teachers need to understand that middle school students are a diverse population and that they will benefit from educational programs that offer variety in content material and presentation style, and that stay current with popular culture (Hagood, 2000). An understanding and application of the basic principles underlying individual learning styles and multiple forms of intelligence will support teachers successful facilitation of middle-level students literacy learning. Teachers' modifications for these individual difficulties may be appropriately created through integrated assignments involving media and non-print materials (Kamil, Intrator, & Kim, 2000, Luke, 2002).

The need for practical applications of literacy have been shown to formalize new learning and to help students understand the importance of their knowledge of reading and writing. The extension of classroom reading projects into community involvement and service learning take literacy beyond the classroom and provide needed connections for young adolescents (NMSA, 2003; Rycik & Irvin, 2004; Erickson, 2003). This survey indicated that many middle schools were not using this important extension of classroom-based reading activities. School organizers, in general, and literacy specialists specifically need to consider approaches to reaching out to local businesses and service organizations to create mentoring in authentic contexts and to build partnerships for learning. Another outlet for providing these connections is by parents and community members visit classrooms and talk to students about needs and opportunities in the community and to have student research initiated in collaboration with community members.

Postsecondary teacher education programs need to pay particular attention to survey overviews such as this one. Knowing that we must consider that a percentage of our

students are assigned middle-level practicum placements for experiential knowledge in the local school districts, finding diverse approaches to teaching reading to this unique age group is important. Our knowledge and preparation of students for these experiences are limited by what the schools offer to readers of all levels. This new information gleaned from the survey will be important for the undergraduate classes and for the graduate programs because it extends knowledge in the area of middle-level research.

This survey project continues the research investigation of middle-level reading for professional and teaching applications by providing an analysis of survey results focused on reading programs (nature and organization of assessment and instruction) and middle-level educational elements. As educators in the state investigate the nature of existing programs for middle-level educational programs, the results help them consider how we can align our programs with the research mandates for more focused programming that meet local, state, and national standards. For example, a more balanced approach to literacy instruction including reading and writing for word identification, comprehension (including vocabulary), fluency need to be modeled and practiced in higher education settings so that this integrated format is second nature to potential reading specialists and literacy program directors. Educators need to extend their understanding or the nature and importance of out-of-class reading (much of this in the form of technology) to include as connections in literacy instruction. The survey results provide practical information that teachers will use as topics for reflection as those in elementary and secondary programs consider the transitions necessary for middle-level students to develop literacy, and collaborative opportunities.

References

Allen, J. (2000). *Yellow brick roads: Shared and guided paths to independent reading 4-12*. Portland, ME: Stenhouse Publishers.

Atwell, N. (1998). *In the middle: New understandings about writing, reading, and learning*, (2nd ed.). Portsmouth, NH: Heinemann.

Combs, M. (2003). R*eaders and writers in the middle grades*, (2nd ed.). Upper Saddle River, NJ: Merrill.

Corbin, J. & Strauss, A. (2007). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Newbury Park, CA: Sage.

Doda, N. M, & Thompson, S. C. (2002). *Transforming ourselves, transforming schools: Middle school change.* Westerville, OH: National Middle School Association.

Eriockson, L. G. (2003). *Applied literacy in the middle grades: Introducing children to authentic inquiry*. Boston: Pearson Allyn & Bacon.

Felner, R., Jackson, A., Kasak, D., Mulhall, P., Brand, S., & Fowers, N. (1997). The impact of school reform for the middle years. *Phi Delta Kappan*, *78*(7), 528-550.

Fountas, G. S., & Pinnell, I. C. (2001). *Guiding readers and writers grades 3-6: Teaching comprehension, genre, and content literacy.* Portsmouth, DE: Heinemann.

Groves, R. M., Fowler, F. J., Jr., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey methodology*. Hoboken, NJ: Wiley-Interscience.

Hagood, M. C. (2000). New times, new millennium, new literacies. *Reading Research and Instruction*, *39*(4), 311-328.

Hox, J., & de Leeuw, E. (1994). A comparison of nonresponse in mail, telephone, and face-to-face surveys: Applying multilevel modeling to meta-analysis. *Quality and Quantity*, *28*, 329-344.

Jackson, A. W., & Davis, G. A. (2000). *Turning points: Educating adolescents in the 21st century*. New York: Teachers College Press.

Kamil, M. L., Intrator, S. M., & Kim, H. S. (2000). The effects of other technologies on literacy and literacy learning. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds*.)*, *Handbook of reading research*, *Volume 3*. (pp.771-788).

Luke, C. (2002). New literacies in teacher education*, Journal of Adolescent and Adult Literacy*, *43*(5), 424-435.

National Middle School Association. (2003). *This we believe: Successful schools for young adolescents*. Westerfield, OH: National Middle School Association.

National Assessment of Educational Progress. (1999). *Reading report card for the nation and the states.* Washington, DC: U. S. Department of Education.

Raphael, T. E., Kehus, M., & Damphousser, K. (2001). *Book club for middle school*. Lawrence, MA: Small Planet Communications.

Redman, P. (1995). Finding a balance: Literature circles and teaching reading. In B. C. Hill, N. J. Johnson, & K. L. Noe (Eds.), *Literature circles and response* (pp. 55-70). Norwood, MA: Christopher Gordon Publishers.

Ruiz-de-Velasko, J., & Fix, M. (2000). Overlooked and underserved: Immigrant students in the U.S. secondary schools. Washington, DC: Urban Institute.

Rycik, J. A., & Irwin, J. L. (2004). Teaching reading in the middle grades: Understanding and supporting literacy development. Boston: Allyn & Bacon.

Unrau, N. (2004). *Content area reading and writing: fostering literacies in middle and high school cultures*. Upper Saddle River, NJ: Pearson / Merrill Prentice Hall.

Francine Falk-Ross is an Associate Professor at Pace University. Her research interests are in content area reading, new literacies, and language factors in literacy development. E-mail: ffalkross@pace.edu.

Diving into an Issue and Dragonflying on the Surface: Two Contrastive Discursive Patterns of Class Discussions

Wen Ma
Le Moyne College

Abstract

While a discussion-based instructional approach is often employed by classroom teachers and university professors for teaching English, literature and varied content areas, there are few studies that have focused on how intellectually sophisticated doctoral students participate in class discussions to acquire discipline-specific content and what their discussions look like. This case study examined the discursive features of discussions in a doctoral seminar on qualitative research methods for education. Two types of dialogic engagement were identified: diving into an issue *and* dragonflying on the surface*. These findings may help professors and advanced learners tap the potential of whole class and small group discussions in graduate seminars.*

As participatory discussions (Britton, 1990) are being recognized as an avenue of learning, a growing number of studies have explored the use of class discussions to help students learn literacy, English, literature, science, and other content areas in K-12 settings (Applebee, Langer, Nystrand & Gamoran, 2003; Evans, 2002; Glazier & Seo, 2005; Kong & Pearson, 2003; Wells, 2001). At the postsecondary level, important research also has been done on how discussions impact college students' learning of discipline-specific content (Fishman & McCarthy, 1998; Goldblatt & Smith, 1995; Nowacek, 2007). While discussions are often employed in varied graduate seminars, there are few investigations on how intellectually sophisticated graduate students participate in class discussions, and what their discussions look like (Prior, 1998).

This gap in research prompted me to ponder more mature learners' participation in discussions and the ensuing features of their discussions. Clearly, such research would be useful in extending the current research base on discussion across grade, ability and curriculum levels. Findings from such studies can also help professors better understand advanced learners' discursive practices and thus effectively engage all students in the learning process.

In this case study, I explored the dialogic practices that took place in a doctoral seminar on qualitative research methods for education. Specifically, the study addressed this research question: What are the discursive features of discussions in the doctoral seminar?

Theoretical Framework

Drawing on Vygotsky's (1978, 1987) sociocultural perspective and Bakhtin's (1981, 1986) theorization of dialogism, Applebee (1996), Kozulin, Gindis, Ageyev and Miller (2003), Lee and Smagorinsky (2000), and Wertsch (1991) made a number of theoretical propositions concerning the value of dialogue as a sense-making, problem-solving tool. Dialogues among the learners provide the needed social context for them to wrestle with curricular content by interacting with each other. Moreover, the dialogic interactions draw their attention to the specific content or issues in the class, and develop their understanding and thinking. The exchanges provide open opportunities for each learner to talk out issues of importance, which further expand an individual's thinking through articulating his or her ideas and hearing the alternative views of others. According to Britton's (1990) studies on children's participatory dialogues with their peers, expressive language use offers the avenue to produce, reshape, and extend one's ideas from an often fuzzy and work-in-progress state to a clearer and more meaningful form. In all, dialogic interactions make it possible for students to intellectually engage with the content of learning.

These theoretical underpinnings for learning have encouraged many classroom teachers and university professors to use participatory discussion as an instructional tool in language arts, literature, and varied content areas. For example, at the elementary level, *book clubs* (McMahon & Raphael, 1997) and *literature circles* (Daniels, 2001) incorporated student-led discussion of literary texts into the teaching and learning processes of language arts and varied content areas. At the secondary level, Barnes and Todd (1995) and Miller (2003) highlighted the cross-curriculum, cross-text implications of peer discussions. At the college level, Goldblatt and Smith (1995) examined a group of first-year college students' conceptions of good discussions, and Nowacek (2007) further called for interdisciplinary connections through class discussions.

In addition to the pedagogical value, previous research has explored the discursive features of class discussions. For instance, Cazden (2001) identified the inquire-respond-evaluate (IRE) classroom discourse pattern based on the teacher-student verbal interactions in her combined first, second and third grades. Moreover, Nystrand (1997) advocated a model of substantive discussion through the use of authentic questions and uptake for teaching English and literature at secondary schools. Nevertheless, while there exists a wealth of research literature in the fields of composition, rhetoric and genre studies on how graduate students socialize into discipline-specific discourse community through discussion, writing and reading (e.g., Berkenkotter & Huckin, 1995; Geisler, 1994; Prior, 1998), there are few empirical studies that focus on doctoral students and doctoral-level discussions. Consequently, the discursive features of doctoral seminars are not well understood by the educational community.

Central to the present inquiry is the construct *discussion*. There are different definitions of class discussion. For example, Dillon (1994) defined discussion as "a unique form of group interaction, where people join together in addressing some question of common concern, something they need to understand, appreciate or decide" (p. 5). Brookfield and Preskill (2005) further proposed using discussion as a forum of social engagement and as a way of teaching. Here I use the term to refer to open-ended, course-related discussions as an instructional tool, including whole class discussions led by the professor and small group discussions by students.

Also pertinent to this study is how to analyze discussions. In the research described below, the class community was orchestrated by the professor to wrestle with various aspects of the disciplinary content through articulating each one's own thinking and understanding of the assigned reading, writing, and the research projects, and through hearing others' interpretations of the course material. As a result of the back-and-forth exchanges, they came to acquire the discipline-specific vocabulary, understand the epistemological framework of the qualitative techniques, and learn the procedures for doing qualitative research in education. To analyze the features of these class discussions, I drew on Fairclough's (1995) three-layered discourse analysis scheme. Specifically, I first carefully recorded and transcribed the oral discussions. Then, I organized and interpreted the transcribed data (i.e. the written texts) by adding the contextual and paralinguistic information. Lastly, I described the thematic issues under question holistically. These analytical procedures are further substantiated in the next section.

In summary, research about the discursive features of a doctor seminar may contribute to the existing research literature on class discussions and the teaching practices following a discussion-based instructional approach. Through examining how diverse students interacted dialogically to make sense of disciplinary content at the doctoral level, this research may contribute to a fuller understanding of the interplay of class discussions, advanced learners, disciplinary content, and cultural and linguistic influences.

Methods

The Research Design

This research followed a descriptive case study design (Creswell, 2005; Merriam, 1998). To choose what to describe about the discussions in the "case," I experienced what Jaworski and Coupland (1999) called the weaknesses of discourse analysis and other qualitative research since

> there will always be problems in justifying the selection of materials as research data. It is often difficult to say why a particular stretch of conversation or a particular piece of written text has come under the spotlight of discourse analysis, and why certain of its characteristics are attended to and not others. (p. 36)

As the student talk and interaction speak for themselves, I selected exemplar discussion data, situated within detailed descriptions of class contexts, to show what was discussed and how each discussion went. I feel that concrete discussions can give readers a real taste of the discourse life of the seminar. The thick descriptions also helped me create a more trustworthy research process.

The Setting and the Participants

The research site was a research-intensive state university in the Northeast. The seminar was a research methods class that all doctoral students at the university's graduate school of education were required to take. Its content, ranging from the epistemological foundations, methodological considerations to actual steps needed to conduct a qualitative study, provided rich opportunities for the education students to explore these issues collaboratively. The coursework consisted of reading, writing and discussing activities. The readings focused on concepts and on methods of doing qualitative research in the field of education, ranging from the qualitative paradigm, to research statements, research questions, design, and procedures for data collection and data analysis. Actual fieldwork was also assigned for the students to put what they learned through reading and discussions into practice, including the completion of an interview case study and a final observational field study. Amidst all these learning activities, a variety of discussion formats were employed, including whole class discussion, small group discussion, and online discussion board. The moment-to-moment discussions across the semester offered a social forum to engage the students' thinking and understanding of qualitative research methods for education as a disciplinary discourse.

The seminar had sixteen American and Asian doctoral students and one advanced master's student. Situated within a larger project that compared the participatory learning experiences of a cohort of mainstream American students and Korean international doctoral students, the research described here involved eight of the students in the seminar. These participants' information is summarized in Table 1 (all names are pseudonyms; for more information about participant selection criteria and the instructor's roles, see Ma, 2004).

It needs to be noted that in previous studies, I examined the culturally and linguistically different students' perceptions of discussions (Ma, 2007), and analyzed their participatory learning experience using a variety of qualitative data, including multiple interviews with the participants, their written work and my participant observations (Ma, 2008). The focus of this study was to describe how they interacted dialogically, and what their discussions look like, drawing primarily on the discussion data.

Data Collection and Analysis Procedures

The role I played was participant observer (Spradley, 1979). To collect the discussion data, I sat in every class mainly to listen, observe and take field notes, and engaged occasionally with the learning activities in the class. This way I not only developed a better sense of the course materials, but also established rapport with the students. As I used the

Table 1
Participants' Profiles

Name	Age	Gender	Ethnicity	Major	Doctoral Experience	Master's/When	Undergrad/When
Bettie	23	F	Euro-American	Psychology	Ph.D. 2nd Yr	Psychology	Psychology & Art/2001
Jerry	30	M	Euro-American	Literacy Ed	Ph.D.1st Yr	English Ed/2003	English/1995
Patty	26	F	Euro-American	Science Ed	Ph.D.1st Yr	Science Ed/2002	Biology/1998
Jumi	28	F	Korean	Reading	Ph.D. 2nd Yr	General Ed/2002	German & English Lit./1999
Minsu	27	M	Korean	Early Childhood	Ph.D. 2nd Yr	General Ed/2002	English/1999
Sunhee	27	F	Korean	TESOL	Ph.D.1st Yr	General Ed/2003	English/2000
Taewoo	27	M	Korean	TESOL	Ph.D.1st Yr	General Ed/2003	English/2000
Yoko	26	F	Japanese	EarlyChildhood	Master's	Master's	English/

first two classes to explain the purpose of my research and to complete the consent forms, I started to videotape class discussions from the third week until the end of the semester, totaling approximately forty-eight hours. My videotaping procedures were as follows: before each class, I set up a camcorder in a far corner of the room; when class started, I turned on the camcorder and let it run until the end of the class. The students soon appeared not paying much attention to the camcorder as they became used to it. For better sound quality, a microphone for the camcorder was placed in the center of the tables around which the whole class sat. Three tape recorders were used to audiotape small group discussions.

The discussions were transcribed selectively because of the large amount of data involved, but I made sure that early, middle and final classes were transcribed in order to assess any indications of change in class discussions over time. Consequently, the first two, the middle two, and the last two class discussions were selected to be transcribed. Thus, six classes, or approximately 44% of the thirteen total class meetings, were transcribed.

To analyze the discussion data, I used the constant comparison method (Glaser & Strauss, 1967). Specifically, I put all the transcription into two columns. Then I placed descriptive transcripts on the left-hand side of the page and the corresponding interpretative account on the right-hand side (cf. Barnes & Todd, 1995). Next, I repeatedly read the discussion transcripts on the left-hand, identified important sections, and wrote descriptive codes to denote major thematic categories on the right-hand margins, e.g. *sharing the participant selection steps* or *the write-up experience for the interview project, examining the concept of generalizability*, etc. To triangulate the discussion data, I constantly watched the videotapes, studied the interviews with the participants, and drew on my field notes of details not captured in videotapes (especially during small group discussions). As my analysis proceeded and more codes appeared, I constantly expanded, modified, or regrouped them into major codes, e.g. *sharing procedures for a research project,* or *using revoicing* (McVee & Pearson, 2003) *to engage others*, etc. until the two final themes reported in the following emerged.

Findings

The discussions in the seminar presented two contrasting discursive patterns. Sometimes the exploration lingered on particular issues, going rather deep into each issue and weighing various aspects of the issue. At other times, the topics changed rather fast, and the conversation touched upon the surface of an issue, and then moved on to a new one. I call the first type of dialogic engagement *diving into an issue*, and the second *dragonflying on the surface.*

Diving into an Issue

A salient feature of the whole class discussions is that the discussants often dived into the heart of a few key issues, or interrelated aspects of one important issue, to look at them in more depth or from different directions. Their exploration for each topic was both prolonged and deep. This was particularly prevalent when the instructor facilitated the flow of the discussion or provided explanations or comments on the course concepts or individual research projects.

Most whole class discussions in the seminar signified such discursive characteristics. For instance, in Class 3 when the class was wrestling with the notion of paradigm drift, there was an extensive examination of Guba and Lincoln's (1994) "Competing paradigms in qualitative research." Led by the instructor, the whole class went over a range of issues, such as for whom the article was written, what was a paradigm, paradigmatic distinctions between qualitative and quantitative research, and practical considerations involved in doing qualitative research. The following segment of the discussion shows how the instructor made use of instructional questions to elicit student response and then provided detailed explanations on these points as a way of teaching the disciplinary content:

> **Dr. Jones**: … Now in their discussion of this right here at the beginning [pointing to the article], they start talking about the received view, the conventional view. They use these particular terms when talking about positivism, which may seem like a very minor thing. But why do they even bother

to do that? Why would they say things like the received view, the conventional view? What does that language indicate to us?

Patty: That's the paradigm, is what I was thinking. That's the current paradigm, and anything outside of it is a breaking paradigm, trying to break the mold.

Dr. Jones: That's the current paradigm? If you think about the term received view, received implies that there is someone there giving something, or that there is something actually there for uptake, and that somebody who is, it's being received blind, right? If you think about the received view, they are talking about the larger view within our society, what is considered the conventional. That's the other term they use, conventional view, the acceptable view of scientific research and inquiry. An alternative term that they may have used would be traditional view. But traditional is somewhat problematic in that ethnography has a long established tradition. It doesn't happen to be the most conventional view when it comes to scientific inquiry in education. Part of what has happened in educational studies is that there was a time when it wasn't clear which way things were going to go. But because of the influence of educational psychology, psychologists started doing experimentation in a variety of subjects, and the history I know about is in literacy. But around the turn of the 20th century, there were people who began doing experiments, and then over time what those people began to argue was that if education wanted to gain in stature as a discipline, if people wanted to gain stature for education, part of what they needed to do was to make education a more scientific enterprise. That meant doing studies that were like what scientists do, quantifying, trying to remove the human element from things [omission of 389 words on the prevalence of quantitative research]. If we think of these different particular paradigms, one reason why they matter and why we need to understand them is because they do affect policy decisions, but there are also other reasons to think about. Why do we bother to read this kind of stuff in a course like this? Why should we think about this?.....

Bettie: Even people that aren't in a quantitative course have grown up where quantitative field data is more valid or accepted. We're not, we are not grown up in a vacuum. And I, even though I feel very, very strongly towards qualitative data, I still hear the messages in my brain, in my department where students who are doing quantitative are more supported and excited, and faculty members are there and ready to work. And so I need to read this to remind myself that, that people are thinking about this sort of thing, and I can take myself out of that line of thought. I don't know. That's why I need to read this.

Dr. Jones: I think that's an important point. We are in this culture …… (Class 3)

Here the instructor talked significantly longer (398 words) than Patty and Bettie combined (140 words). A tilted participation pattern is obvious. The whole exchanges began with teacher-directed questions about *the received view*. Based on Patty's brief response to the question, Dr. Jones went on to elaborate on what the received view as a paradigm implies in the current society, the historical context of educational research and the national policy-making regarding the issue of funding scientifically-based research. This question occurred to me as a rhetorical one because Dr. Jones was already prepared to give much more details than Patty asked for. Yet such substantive information appeared useful to help Patty and the other students better understand the larger socio-cultural context for the so-called paradigm wars and their effects on qualitative research, for it was unlikely for the beginning doctoral students to be fully aware of all this vast range of conceptual background for the disciplinary content.

The instructor then asked further questions that guided students' thinking about why they needed to read this piece. With Bettie articulating her thought to justify learning and using qualitative research techniques in her field of study (Psychology), Dr. Jones made further lengthy comment. Afterwards, she posed further questions; based on students' responses, she would weave more information into her informal lecture to facilitate the students' thinking and understanding of the course material.

Another example may be found in the whole class discussion in Class 4, when they explored concepts of qualitative paradigm and epistemology. After the students first engaged in small group discussions, the whole class got together to consider Cunningham and Fitzgerald's (1996) "Epistemology and reading." Again, the instructor used a series of questions to propel students' thinking to a deeper level. Notice how Dr. Jones responded to each student's response:

Dr. Jones: …What tips you off to what somebody's paradigm or epistemology is, if they don't actually tell you because not every piece of research that you read, in fact probably most, will not tell you right up front. They won't say, "Okay, I've been reading, you know, Cunningham and Fitzgerald, and I'm writing, so you know, from positivist perspective or I'm writing from a hypothetical deductive, so this perspective or I'm writing for this." They're not gonna tell you that straight off the bat. So how do you know where people are coming from when you analyze research?

Bettie: If they present ideas and then at the very end they say, um we as researchers want you to think of your own ideas, we want you take this as information but we want you to test this out, don't take our word for it and then that's a clue because they want you to a they're not offering you the knowledge and expecting you to take it up as knowledge.

Dr. Jones: What are they offering?

Bettie: They're offering a real possibility to the knowledge.

Dr. Jones: And what does that imply?

Bettie: That there're other possibilities. That you have your own possibilities, but they could be wrong, or, or whatever, from that framework. They, they're implying that, that knowledge could, could possibly not be what you need, that there are other options.

Dr. Jones: They don't have to. That person is not in the position of having to say "this in the end is the proof that this is true." Although people, people have talked quite that way in articles, but essentially you can see, sometimes that people in the end are basically saying "and this is the proof that what I started out with is true, and this is truth." So, one way is to look at how they situate it that way. What else do you look for? What tips you off in terms of somebody's position? Or you could be just listening to somebody talk, you could go to a research presentation, or you're listening to one of your professors in class, um what tips you off to what somebody's research is?

Taewoo: Actually, I think epistemological beliefs really affected what kind of research the researchers are trying to do. I mean the type of research. I thought if they have like higher epistemological belief, I think they try to use qualitative research. On the other hand, if they're like quantitative type of person, I think he may, like, haven't a lot of difference on the lower level of epistemological belief. So, I think the type of research gave us a tip whether, what kind of epistemological beliefs he has.

Dr. Jones: So the types of questions the person asks, the types of research, the research focus and the research questions also tip you off to what paradigm he might be looking at it and what epistemology he has. I think you are definitely right. You don't want it, you don't want to fall into a trap of saying "higher and lower epistemology" as if there's an hierarchy of qualitative, naturalistic inquiry, as if they're talking "we're the best, and all positivist are on the bottom." They're down here (pointing to a chart), doing lower level thinking. We don't. Those types of positions are not helpful. And in certain cases you would be greeted with a severe backlash, or people who were greatly offended. But we want to avoid those types of refutations. But if you look at questions that people ask and the research focus, that will also tip you off to how they're positioned …… (Class 4)

In this part, Dr. Jones again spoke (391 words) much more than the students did (206 words). What should be noted is that she used *revoicing* (McVee & Pearson, 2003) to restate, summarize, or paraphrase words or ideas expressed by the students for the purpose of expanding, clarifying, or summarizing information to further their thinking and understanding. Using this method, Dr. Jones twice pushed Bettie's articulation, hence thinking, with her two non-rhetorical questions "What were they offering?" and "And what does that imply?" Dr. Jones also followed up part of Taewoo's response to guide his understanding about how to recognize a researcher's epistemological stance, without ever directly saying that Taewoo was wrong in thinking about "high and low epistemological beliefs." Student talk thus helped Dr. Jones know their level of understanding of the course content and then became teachable moments for her to provide the needed information to extend their learning.

Although not all students participated in either of the two particular segments, the instructor's explanations contained substantive information. Because such critical information was directed towards the whole class, all students' learning and understanding were scaffolded collectively toward what Bruner (1986) referred to as the mentally "higher ground" (p. 73). At any point, a question may come up from the class, and the instructor's "mini-lecture" turned into a kind of open discussion. Importantly, this discursive pattern appeared repeatedly during the instructor-student interactions in many classes across the semester.

Dragonflying on the Surface

Sometimes, especially during small group discussions when the students were instructed to synthesize key ideas of an article, to discuss methodological issues related to the design of a study in the assigned reading, or to share problems and experiences based on their own studies, the topics and the pace of exchanges were left under the control of individual groups. At these times, the students often touched upon many issues or topics in passing, but they did not long delve into any particular one(s) for an in-depth examination, resembling dragonflies touching the water on the surface and then swiftly flying onto the next spot for new prey.

An example this type of dragonflying may be found in a small group discussion in Class 12, when students were asked to share their case study in small groups. In one of the groups composed of one American student (Jerry) and two Korean students (Taewoo and Minsu), after Jerry introduced his study, Taewoo and Minsu asked a range of questions, each followed by some response from Jerry, and then the discussion moved to a new topic, as shown below:

Taewoo: So, you're doing your writing?

Jerry: Yeah, I've started to write a little bit. The hard part for me is to combine the theories that I'm working on. The positioning theory and what's called critical discourse analysis. When I look at the way the students are talking and the way that they're using their speech, and the way they talk to one another, to position themselves in the class, it's difficult to put the two together. But I don't think I can do it any other way. Because they're positioning themselves through their discourse, I

need both. I can't just rely on one. So that's, that's the hard part.

Taewoo: So, your topic for your research is like—

Jerry: It's how physically disabled boys position themselves in a classroom, in a remedial reading classroom.

Taewoo: So, you're going to use, like that narrative way of writing?

Jerry: Yeah, I'm going to present it in a narrative form. But that's why I asked you two earlier about the literature. Because I have so much, I don't want to spend, y'know, 10 pages, just reviewing literature. I wanta, like she said, find a way to kind of intertwine it in as best I can. But I don't wanta do that in an incorrect way, you know. So, that's probably the biggest problem I have, rather than the theories. Is putting the literature in the right places, y'know, because I have literature on disability and positional theory on critical discourse on Native American learning styles and, [cough] excuse me, on all kinds of other things that are all relevant to what I'm doing.

Taewoo: So you're going to use all of it, or just some?

Jerry: Well, I use bits and pieces, I think, of all of that because it all supports what I'm trying to say. So, but the problem, another fear that I have is that it may, I don't want it to look like I'm taking on Ricky [one of the participants in Jerry's case study]. Because by, maybe I won't even include it, that stuff, because I'm pointing out one different student. I don't know if I should.

Minsu: So you're, you're writing the, all the 5 students, right? You're writing about all?

Jerry: Yeah. I think I have to, unless, I don't know, I guess I do have to, you know.

Taewoo: It's a really looong story.

Minsu: Comparing my work, yours is really tough.

Jerry: Yeah, well, it's really good for me. It takes good writing skills.

Taewoo: I have to make my progress eventually. That's why I'm doing this free research right now. It's terribly hard to focus on research subject.

Jerry: Yeah, I think mine might be longer, but I think it's easier to do five people instead of just one. That's gotta be really hard. Not to put any pressure to you. (laugh)

Taewoo: I feel I like to change my topic sometimes.

Jerry: It's a little late for that. (laugh)

(Silence for 6 seconds)

Taewoo: What time does the class end?

Jerry: 6:50. (Class 12)

The seven questions Taewoo and Minsu asked included whether Jerry had started writing, the format of his writing, and whether he would include all five participants; each question was followed by Jerry's comments or explanations. Then the discussion turned to the difficulty in Taewoo's writing and the challenges he faced, ending in their checking the time, waiting for the class to end. The feedback did not seem to have been focused on a careful examination of Jerry's study, and reflected little personally invested, serious wrestling with the issues in his study.

A similar pattern can be found in many small group discussions. As a participant observer, I feel that intense follow-up scrutiny and inquiry-oriented engagement were almost unexpected. Still, during interviews and informational conversations with me, the participants maintained that such small group exchanges were useful for each individual to talk about his or her individual appropriation of the common course content in a more democratic and less threatening atmosphere. My own observations confirmed that during small group discussions there was no lack of interest in and support for each other's studies, as demonstrated by frequent laughing, eye contact while talking or listening, clarifying questions or short comments going back and forth. However, sometimes the quality of the feedback appeared not substantive enough to push forward the original inquiry. As a result, discussions of one another's studies sometimes became mini-presentations, more for sharing than for critiquing each other's work.

Discussions that remained at literal comprehension level rather than leading to deeper exploration of the issues under question may be viewed as a special type of dragonflying. For example, during a ten-minute small group discussion of Alvermann, O'Brien and Dillon's (1996) "On writing qualitative research" in Class 13, the discourse of Jumi, Sunhee and Yoko's group (one of the rare occasions for an all-Asian discussion group), while querying the notions of *writing* and *write-up* as described in the article, showed a dragonflying-type of engagement. Take a look at their discussion transcript.

Yoko: So, what is the difference between writing and write-up? …

Sunhee: I always thought that they are the same one, but it seems like they actually differentiate writing and write-up.

Yoko: The last part.

Sunhee: So, write-up is for the final project?

Jumi: That's what I understand. David said the writing before write-up.

Yoko: Write-up focuses on final representation. So representation …

Sunhee: The author is saying that we have to keep writing.

Yoko: But we have to write up to complete the project.

Sunhee: How can I see (inaudible) put that notion in my own research, to write up, keep writing, and then another process for my write-up, for my final thing.

(Silence for 10 seconds)

Yoko: Writing, writing is a form of, Wolcott says, writing is a form of formal thinking. So when you are writing, you are representing your own thoughts, what [is] in your mind, everything could be possible. But after I went through reading all the writing that you have written, then like we say "We need to integrate (inaudible) pieces of information.

Sunhee: (inaudible) I was writing and then in spite of, I have information

Yoko: (interrupts) Yeah, Glesne.

Sunhee: Everything comes together.

Jumi: I think it means keep writing, you know. Keep writing. (laugh)

Yoko: I'm doing best (inaudible)

(Silence for 6 seconds)

Jumi: It says, which data I should put in my final article, and which data shouldn't I use.

Yoko: I didn't start my writing, so I don't know, but it is difficult for me. This data will be useful. We have been updated for (inaudible) limitation is 15 to 20 pages, right?

Jumi: Yeah, we have to select a good one. (Class 13)

There were several noticeable discursive features in this part. First, in this relatively short segment of discussion, there were frequent pauses (three pauses of 6 seconds or longer). This suggests that the discussion probably did not proceed smoothly, or the discussants were perhaps having some difficulty understanding the notions of writing and write-up as distinguished by Alvermann et al. (1999). Although the places where they got stuck may be areas where they were involved in quiet thinking, and there were discourse traces showing that these participants did try to make some inter-textual connection between Alvermann et al.'s (1996) article and Glesne's (1998) book, their collective efforts did not move much beyond trying to figure out the literal meaning of the two terms. In addition, none of the discussant's utterances was long or extensive, which, as the content of their utterances showed, made it hard to articulate a substantive idea in one turn. Importantly, while their discussions centered on the differences between writing and write-up, they did not appear to be clearly aware of the idea suggested in the article, which was that the writing process, as a way to engage with and think about the data, prepares for, shapes, and sharpens the final form of the write-up.

These points suggest that the Korean and Japanese students were probably more concerned with the finished product than taking the recursive process to grapple with the data for a richer and more thoughtful written product (Wolcott, 2001). Still, the interrogations appeared to help them move toward understanding the issues under question despite that none of them went deep enough individually, nor did they probe extensively as a group.

Conclusion

The above discussion snapshots represent two discursive patterns as this cohort of doctoral students grappled with qualitative research methods dialogically. While facilitating the whole class discussions, Dr. Jones talked lengthily and directed the students to go rather deep into key course concepts, methodological procedures or individual projects. Like plunging a boulder in the water, the instructor led the class to dive into the heart of these tasks. During student-led small group discussions, however, the conversational topics changed rather fast, and the exploration often touched upon an issue and then quickly moved onto another one. Like skipping a rock on the water, the discussants dragonflied on the surface of numerous topics.

These findings underscore the irreplaceable, yet reciprocal responsibility of the learned and the learner in a discussion context, even at the doctoral level. As evidenced in the whole class discussion about the received view in Class 3, the class discourse community led by Dr. Jones came to intellectually gaze at the notion, weighing the additional background information and the on-the-spot facilitation, cross-examining it from diverse perspectives, and thus able to arrive at a deeper understanding of the notion. Similar patterns can be observed in Class 4 and other classes as well. At the same time, the small group discussions across the semester afforded more peer interaction time, when different learners (e.g., Jerry, Minsu and Taewoo in Class 12, or Jumi, Sunhee and Yoko in Class 13) shared their developing thought or work to shape up their grasp of the disciplinary content, conveniently and comfortably. In this sense, diving into the heart of an issue and dragonflying on the surface of multiple issues relevant to the student's thinking and understanding of the disciplinary content are both meaningful discussions for learning: the former for the depth of exploration, and the latter for the breadth of topics.

These findings may have both theoretical and pedagogical implications. Although this study focused on the discursive features of the seminar holistically, the data show that individual students may have different participation and response preferences. For example, during both whole class and small group discussions, some students (Bettie and Patty in Class 3, or Jerry in Class 12, for example) talked much more than their counterparts did, and they appeared adept at using discussions to develop their thinking and learning. This reconfirmed what Britton (1990) long theorized as the value of student talk: articulation helps to shape and clarify thoughts.

Nevertheless, I noticed different participation and response styles in the same seminar. For instance, two of the ten (20%) mainstream American students and all seven Asian international students appeared reticent during discussions,

who admitted during interviews and informal conversations with me that they felt pressured to participate in the dialogic exchanges. Consequently, they listened much more than they talked, as reflected through the small group discussion examples in Class 12 and Class 13. Being a participant observer, I feel this difference may be related to a student's literacy skills, trying to manage the complex disciplinary content in fast-paced discussions, but may also have to do with one's prior educational experiences for appropriate ways of participation, by having a different cultural upbringing, and by one's unfamiliarity with class discussion (cf. Liu, 2001; Ma; 2007; Watkins & Biggs, 1996).

If discussion is not be a neutral avenue of learning, to engage diverse learners in the learning process, the professor of a discussion-based graduate seminar may tap the dialogic potential by alternating whole class and small group discussions in order to maximize opportunities for all students to participate. The professor of a multicultural class may also change the dominant or subordinate subcultures by promoting mutual understanding and reciprocal accommodation among the students. Besides small group and whole class discussions, the professor can give pre-discussion reviews of the content and post-discussion summaries of the discussions to better prepare the students for the discussions. To further extend such discussions from more advanced students to younger ones, educators in diverse settings may employ a variety of instructional activities and strategies that aim to cultivate all learners' participation incrementally across their academic ability levels.

Additionally, these findings may help professors recognize the complexity and diversity of class discussions in different seminars as similar discursive practices seem transposable into other seminars, whether it is literary theory or ancient philosophy. Therefore, with a better understanding of how doctoral students participate in discussion and what their discussions look like, professors of different seminars can proceed strategically in ways that connect the specific disciplinary content with varied types of discussions, flexibly combining them to maximize the full potential of discussions (Ma, 2008; Prior, 1998).

Finally, while this case study provides important information about the discursive features of a doctoral seminar, some limitations must be acknowledged. Firstly, the findings only involved eight participants across four classes. More discussion data and more participants should be included to examine and understand the complex relationships among discussion, learning, culture, and disciplinary content. Secondly, it is not very clear whether the discursive features identified in this study are unique to doctoral students. It will be meaningful to study the discursive practices across a broader spectrum of academic ability. Thirdly, as a teacher's scaffolding and facilitation obviously affect students' discussion and learning, the professor's role in graduate seminars needs critical scrutiny (cf. Dillon, 1994; Brookfield & Preskill, 2005). All these issues require further exploration in future research.

Acknowledgements

I thank the editors and the anonymous reviewers for their comments and suggestions. I give my special thanks to all the participants in this study. I also thank Suzanne Miller and Mary McVee for their guidance in the development of this article.

References

Alvermann, D. E., O'Brien, D. G., & Dillon, D. R. (1996). On writing qualitative research. *Reading Research Quarterly, 31*(1), 114-120.

Applebee, A. N. (1996). *Curriculum as conversation: Transforming traditions of teaching and learning*. Chicago: University of Chicago Press.

Applebee, A. N., Langer, J. A., Nystrand, M., & Gamoran, A. (2003). Discussion-based approaches to developing understanding: Classroom and student performance in middle and high school English. *American Education Research Journal, 40*(3), 685-730.

Bakhtin, M. M. (1981). *The dialogic imagination: Four essays by M. M. Bakhtin*. Austin, TX: University of Texas Press.

Bakhtin, M. M. (1986). *Speech genres and other late essays*. Austin, TX: University of Texas Press.

Barnes, D., & Todd, F. (1995). *Communication and learning revisited: Making meaning through talk* (rev. ed.). Portsmouth, NH: Boynton/Cook Publishers.

Berkemkotter, C., & Huckin, T. N. (1995). *Genre knowledge in disciplinary communication: Cognition/culture/power.* Hillsdale, NJ: Lawrence Erlbaum.

Britton, J. (1990). Talking to learn. In D. Barnes, J. Britton, & M. Torbe (Eds.), *Language, the learner and the school* (4th ed., pp. 89-130). Portsmouth, NH: Heinemann.

Brookfield, S., & Preskill, S. (2005). *Discussion as a way of teaching: Tools and techniques for a democratic classroom* (2nd ed.). San Francisco: Jossey-Bass Publishers.

Bruner, J. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.

Cazden, B. C. (2001). *Classroom discourse: The language of teaching and learning* (2nd ed.). Portsmouth, NH: Heinemann.

Creswell, J. W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (2nd ed.). Upper Saddle River, NJ: Pearson.

Cunningham, J. W., & Fitzgerald, J. (1996). Epistemology and reading. *Reading Research Quarterly*, *31*(1), 36-60.

Daniels, H. (2001). *Literature circles: Voice and choice in book clubs and reading groups* (2nd ed.). York, ME: Stenhouse Publishers.

Dillon, J. T. (1994). *Using discussion in classrooms*. Buckingham, UK: Open University Press.

Evans, K. (2002). Fifth-grade students' perceptions of how they experience literature discussion groups. *Reading Research Quarterly, 37*(1), 46-69.

Fairclough, N. (1995). *Critical discourse analysis*: *The critical study of language*. London: Longman.

Fishman, S. M., & McCarthy, L. (1998). *John Dewey and the challenge of classroom* practice. New York: Teachers College Press.

Geisler, C. (1994). *Academic literacy and the nature of expertise: Relations between reading, writing, and knowing in academic philosophy*. Hillsdale, NY: Lawrence Erlbaum.

Glaser, B., & Strass, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Glazier, J., & Seo, J. A. (2005). Multicultural literature and discussion as mirror and window? *Journal of Adolescent and Adult Literacy, 48,* 686-700.

Glesne, C. (1998). *Becoming qualitative researchers: An introduction* (2nd ed.). Boston: Allyn and Bacon.

Goldblatt, E., & Smith, M. W. (1995). Alone with each other: Conceptions of discussion in one college classroom community. *Linguistics and Education, 7*, 327-348.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). Thousand Oaks, CA : Sage.

Jaworski, A., & Coupland, N. (1999). Introduction: Perspectives on discourse analysis. In A. Jaworski, & N. Coupland (Eds.), *The discourse reader* (pp. 1-44). London: Routledge.

Kong, A., & Pearson, P. D. (2003). The road to participation: The construction of a literacy practice in a learning community of linguistically diverse learners. *Research in the Teaching of English, 38*(1), 85-124.

Kozulin, A., Gindis, B., Ageyev, V., & Miller, S. (Eds.). (2003). *Vygotsky's educational theory in cultural context*. Cambridge, UK: Cambridge University Press.

Lee, C. D., & Smagorinsky, P. (Eds.). (2000). *Vygotskian perspectives on literacy research: Constructing meaning through collaborative inquiry*. Cambridge, UK: Cambridge University Press.

Liu, J. (2001). *Asian students' classroom communication patterns in U.S. universities*. Norwood, NJ: Ablex.

Ma, W. (2004). *Intellectual gazing: Participatory learning in a graduate seminar*. Unpublished doctoral dissertation, University at Buffalo, Buffalo, NY.

Ma, W. (2007). Dialoging internally: Participatory learning in a graduate seminar. In C. C. Park, R. Endo, S. J. Lee, & X. L. Rong (Eds.), *Asian American education: Acculturation, literacy development, and learning* (pp. 167-195). Charlotte, NC: Information Age Publishing.

Ma, W. (2008). Participatory dialogue and participatory learning in a discussion-based graduate seminar. *Journal of Literacy Research, 40*(2), 220-249.

McMahon, S., & Raphael, T. (Eds.). (1997). *The book club connection: Literacy learning and classroom talk*. New York: Teachers College Press.

McVee, M., & Pearson, P. D. (2003). Talking the talk: A close examination of teacher-student discourse around written artifacts. *Literacy Teaching and Learning, 8*(1), 47-71.

Merriam, S. B. (1998). *Qualitative research and case study applications in education* (rev. & exp. ed). San Francisco: Jossey-Bass Publishers.

Miller, S. (2003). How literature discussion shapes thinking: ZPDs for teaching/learning habits of the heart and mind. In A. Kozulin, B. Gindis, V. Ageyev, & S. Miller (Eds.), *Vygotsky's educational theory in cultural context* (pp. 289-316). Cambridge, UK: Cambridge University Press.

Nowacek, R. S. (2007). Toward a theory of interdisciplinary connections: A classroom study of talk and text. *Research in the Teaching of English, 41*(4), 368-401.

Nystrand, M. (with Gamoran, A., Kachur, R., & Prendergast, C.). (1997). *Opening dialogue: Understanding the dynamics of language and learning in the English classroom*. New York: Teachers College Press.

Prior, P. A. (1998). *Writing/disciplinarity: A sociohistoric account of literate activity in the academy.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Spradley, J. P. (1979). *The ethnographic interview.* New York: Holt, Rinehart and Winston.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1987). Thinking and speech. In R. Rieber and A. Carton (Eds.), *L. S. Vygotsky, Collected works* (Vol. 1; N. Minick, Trans.; pp. 39-285). New York: Plenum.

Wells, G. (Ed.). (2001). *Action, talk and text: Learning and teaching through inquiry*. New York: Teachers College Press.

Watkins, D. A., & Biggs, J. B. (Eds.). (1996). *The Chinese learner: Cultural, psychological, and contextual influences*. Hong Kong: The University of Hong Kong Press.

Wertsch, J. V. (1991). *Voices of the mind: A sociocultural approach to mediated activity*. Cambridge, MA: Harvard University Press.

Wolcott, H. F. (2001). *Writing up qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.

Wen Ma is an Associate Professor at Le Moyne College. His current research interests include the use of discussion as an instructional tool, multicultural teaching strategies, and immigrant adolescents' learning and socialization experiences. E-mail: maw@lemoyne.edu.

Principals and School Factors That Impact Elementary School Student Achievement

Sharon Gieselmann
University of Evansville

Abstract

This study examined principals and school factors associated with elementary school student achievement. Nine predictor variables were analyzed to determine their impact on student state assessment scores: (a) years of principal experience, (b) years of teaching experience by the principal, (c) years of principal experience at present site, (d) highest level of education by the principal, (e) principal gender, (f) principal leadership as measured by the three subscales of the Principal Instructional Management Rating Scale, and (g) free/reduced lunch population at the school. Study findings are discussed.

Conceptual Framework

During the present school accountability era, identifying specific factors that help schools steadily raise the level of student achievement is important to educational stakeholders including school administrators, teachers, parents, and policymakers. Several large-scale literature reviews have been conducted and found associations between factors related to principals and their schools that contribute to increased student achievement (Leithwood, Begley, and Cousins, 1990; Hallinger and Heck, 1996). For example, one body of research makes the conceptual linkage between factors such as school leadership and the school improvement process. Fullan (2002) indicates that principals who are prepared to handle complex, rapidly changing environments can execute reform efforts that lead to sustained improvement in student achievement. However, another body of research determined that school factors such as socioeconomic status impacted student achievement. Schools with predominantly lower socioeconomic status students are likely to perform less well academically than their counterparts who have mixed socioeconomic status populations (Sirin, 2005).

Are there specific factors that impact student achievement? The literature describes several common elements that appear to correlate positively with student achievement, hence their selection for this study. The rational for their inclusion in this research was that if earlier studies produced significant results, these factors should be examined to determine their impact on student achievement in this particular state. Specifically, nine factors were analyzed to determine their impact on student state assessment scores: (a) years of principal experience, (b) years of teaching experience by the principal, (c) years of principal experience at present site, (d) highest level of education by the principal, (e) principal gender, (f) principal leadership as measured by the *three* subscales of the *Principal Instructional Management Rating Scale,* and (g) free/reduced lunch population at the school. Findings from the literature are included in subsequent paragraphs.

Bista and Glasman (1998) discovered a positive relationship between total *years of principal experience* and school improvement. The researchers discovered that total years of principal experience equated with more effective leadership abilities that impacted student achievement. Hallinger and Murphy (1985) determined that principals serving longer, were able to focus on accomplishing the school's mission while Young (1993) found that principals were more collaborative with decision making as their years in school administration increased.

Grady and O'Connell (1993) reported that *principals with more teaching experience* implied better preparation and understanding of school administrative functions. Additionally, the assumption that increased teaching experience equates with better preparation for the curricular and accountability demands found in administration emerged from a study completed by Shakeshaft (1989). Bista and Glasman (1998) also reported that the most important predictor of leadership ability was teaching experience. Teaching experience may prepare administrators for the varied, day-to-day operations of the principalship.

Young (1993) indicated that *years of principal experience at the current school site* impacted student achievement. Principals serving for more than two years at their present site engaged their staffs in collaborative curriculum development, a precursor to improved student learning, much more than those administrators who were new to their leadership position. Other researchers determined that principals who worked for more years at their present site were more apt to collaborate with teachers to improve instruction and formulate a shared vision that organized all elements of the school around increasing student achievement (Bista and Glasman, 1998).

Grady and O'Connell (1993) reported that *principals with higher levels of education* implied better preparation and understanding of school administrative functions. Hallinger and Murphy (1985) described the formal education experience of the principal as a correlate with student achievement. Additionally, several studies indicated that increased levels of

education helped school leaders apply professional leadership abilities that impacted student achievement (Gross & Herriott, 1965; Hemphill, Griffiths, & Frederiksen, 1962).

In their meta analysis of the literature that related to *principal gender* and student achievement, Shakeshaft, Brown, Grayson, Brunner, Grogan, and Hackney (2006) reported that the *gender of the principal* had impacted student achievement in several studies. Furthermore, many researchers claim that principal gender relates to the exertion of different leadership strengths in school administration such as collaborative leadership, which relates to increased student achievement (Eagly, Karau, & Johnson, 1992; Kochan, Spencer & Mathews, 2000; Shakeshaft, 1989).

Meta analyses of school effectiveness literature (Hallinger & Heck, 1996; Marzano, Waters, & McNultry, 2005) concluded that effective principals exerted influence on school processes directly linked to student learning. Achieving high standards in classroom practices during reform requires sound leadership from school principals, recognized as key players for school success as they supervise and organize the work of others (Waters & Grubb, 2004). Furthermore, reform efforts may be short-term and superficial without strong leadership characterized by instructional capacity building in a cohesive professional community (Spillane & Thompson, 1997).

In addition to these factors exhibited by principals that are related to school improvement, other research describes the school factor of *free/reduced lunch population* as indicative of school achievement. Bista and Glasman (1998) found that schools with predominantly low socioeconomic status students were more likely to have lower student achievement. More recently, Sirin (2005) acknowledged that with all things equal in our schools, as student socioeconomic status increases, so does student achievement. Other research supports the connection between student achievement and free/reduced lunch populations as well (Slovacek, Kunnan, & Kim, 2002; Bulach, Malone, & Castleman 1995).

In sum, during the present reform cycle when schools are under increased pressure to meet yearly state and federal assessment goals, it is relevant to determine if these school and principal factors supported by the literature continue to predict increased student achievement in this particular state. The next section provides study methods and procedures for sample selection.

Study Methods and Sample

Sample Selection

The researcher chose elementary school principals in one state heavily involved with comprehensive school reform efforts since the early 1990s for study participation. This state's education reform act altered the school principal's instructional leadership role significantly when high-stakes testing was implemented in the early 1990s and schools were held accountable for student achievement. Additionally, because the elementary school is organizationally less complex than the secondary school, assessing elementary principal leadership skills may be easier. Bista and Glasman (1998) stated that elementary principals are more likely to affect student performance more forcefully and effectively than administrators at the secondary level. Further, Young (1993) determined that principals serving for more than two years at their present site engaged their staffs in collaborative curriculum development, a precursor to improved student learning. Therefore, elementary principals having served a minimum of three years were selected to participate in the study as they presumably had applied leadership skills that impacted the educational environments at their school sites and student performance. State school directories provided names, addresses, and telephone numbers of elementary principals serving at their present site for a minimum of three years. Principals meeting these criteria totaled 340 and comprised nearly equal numbers from rural, suburban, and urban schools throughout the state.

This study used multiple regression, a non-experimental statistical approach, and addressed the question: What principal and school factors predicted student achievement as measured by the state assessment? In the multiple regression, the predictor variables included (a) highest level of education obtained by the principal, (b) years of principal experience, (c) years of teaching experience by the principal, (d) years of principal experience at present site, (e) principal gender, (f) principal leadership determined by the *Principal Instructional Management Rating Scale* (Hallinger, 1985), and (g) free/reduced lunch population at the school.

The criterion variable was elementary school student achievement measured by the state assessment. A school's comprehensive score on the state assessment was matched with the particular principal participating in this study. This information was obtained from the state's department of education. At the elementary school level, the state assessment included a national norm-referenced test, the *Comprehensive Test of Basic Skills* (*CTBS/5 Survey Edition*), and a standards-based test that specifically measured student progress on state content standards using a multiple choice and open-response writing format. These tests were administered during a two-week testing window during the spring semester. Sample size for the multiple regression analysis was determined by recommendations from Stevens (1996) indicating 15 participants per predictor variable. The study included nine predictor variables; therefore at least 135 respondents were needed for the study.

Principal Leadership Instrumentation

The *Principal Instructional Management Rating Scale (PIMRS)* was utilized for this study because it has been viewed as "the most commonly used instrument in studies that employed an instructional leadership perspective" (Hallinger & Taraseina, 2001). Hallinger described the instrument as "useful for school evaluation, staff development, research, and district policy analysis" (p. 2). The *PIMRS* contains

three dimensions of instructional leadership: (a) *Defining the School's Mission*, (b) *Managing the Instructional Program*, and (c) *Promoting a Positive School Learning Climate*, which highlight leadership functions necessary for this reform cycle. The three dimensions are further separated into 10 subscales that contain a total of 50 items for principal response. The first dimension of instructional leadership contains two subscales, *framing the school goals* and *communicating the school goals*. This subscale contains 10 items that determine if the principal has a clear mission focused on the academic progress of students and whether or not they communicate that mission widely to the school community (Hallinger & Taraseina, 2001).

The second dimension of the *PIMRS* is *Managing the Instructional Program*. This encompasses three leadership subscales: *supervising and evaluating instruction, coordinating the curriculum, and monitoring student progress*. These 15 items assume that even in larger schools, a key leadership responsibility of the principal is developing the school academic core.

The third dimension of the *PIMRS* is Promoting a Positive School Learning Climate and includes five subscales: protecting instructional time, promoting professional development, maintaining high visibility, providing incentives for teachers, and providing incentives for learning. These 25 items are broader in scope and intent and describe successful schools as those creating an "academic press" by developing high standards and expectations along with a culture of continuous improvement. The authors report acceptable validity and reliability data. All 50 items use a 5-point, Likert-type scale: (1) Almost Never to (5) Almost Always. Principals who obtain a high rating on one of the leadership subscales are perceived as engaged more frequently in instructional leadership practices and behaviors associated with principals in effective schools (Hallinger & Taraseina, 2001).

Hallinger (1985) determined that the *PIMRS* met appropriate validity and reliability measures. For instance, when determining content validity of the *PIMRS*, Hallinger (1985) asked school administrators to assign potential items from a randomly ordered list into 10 leadership subscales. The remaining 50 items received at least 80% inter-rater agreement, which Latham and Wesley (1981) considered acceptable. Hallinger (1985) also established construct validity by examining school documents related to instructional leadership and found they described a principal's leadership similar to that obtained from the *PIMRS*.

Last, Hallinger and Murphy (1985) reported internal consistency reliability coefficients for the subscales scores obtained from the *PIMRS*. All subscales were at least .80 using Cronbach's test of internal consistency, which is acceptable according to Latham and Wesley (1981)

Data Collection

All 340 principals were sent a mail-out survey containing a *Principal Biographical Data Sheet* to acquire principal gender, principal age, years of experience as a principal, years of teaching experience, highest level of education, and free/reduced lunch population at the school. Principals also indicated total years of principal experience at their present sites to confirm data reported in the *State school directory*. Principals completed the *Principal Biographical Data Sheet* and self-reported information regarding the variables listed above.

Principals were also asked to complete the *PIMRS* instrument to determine their leadership skills on the three subscales of this instrument: *Defining the School's Mission, Managing the Instructional Program, and Promoting a Positive School Learning Climate.* The researcher utilized the tailored design method outlined by Dillman (2000) for this study: Five contacts were made with respondents: (a) pre-notice letter, (b) survey instrument/consent letter, (c) follow-up post card, (d) replacement letter and survey, and (e) final contact.

Study Limitations

As with all empirical research, this study had certain limitations. First, all participants were from one state and it is possible that individuals from other states might have responded differently to the *PIMRS*. A second limitation was that this study utilized self-reported information based on perception, not actual behaviors. A third limitation was that the criterion variable, student performance on the state assessment, is just one measure of student achievement, not a sole indicator of school achievement. However, while this study had limitations, it expanded the knowledge base about predicting elementary school student achievement using principal and school factors.

Analysis

Principals from 180 of the 340 schools returned survey instruments resulting in a 53% overall response rate. According to Babbie (1990), "A response rate of at least 50 percent is generally considered adequate for analysis and reporting" (p. 182*).* Preliminary analyses consisted of descriptive statistics for the major variables and a reliability analysis of the *PIMRS*. The coefficient alpha for the 50-item composite score was .97. Nunnally (1967) recommended a minimum of .60 for use of a composite score in statistical analysis indicating acceptability of the instrument. The established coefficient alphas for the three subscales were*: (a) Defining the School's Mission (.90), (b) Managing the Instructional Program (.92),* and *(c) Promoting a Positive School Learning Climate (.94).* In this study all subscale scores were at least .80 using Cronbach's test of internal consistency, which is acceptable according to Latham and Wesley (1981).

The researcher utilized multiple regression to examine significant predictors of student achievement. Standard multiple regression where all the predictor variables are entered into the equation simultaneously was conducted to determine the significance of the equation. Multiple regression yielded standardized Beta weights for the significant predictors to

indicate the contributions made by each on elementary school student performance measured by the state assessment. Multiple coefficients of determination (R^2) were computed to determine the relative strength of predictor variables in explaining the percent of variance in student outcomes. The .05 level of significance was used.

The assumptions of independence, normality, and constant variance for multiple regression were checked prior to analysis. Histograms were constructed to assess independence and normality. Residuals indicated a fair approximation to a normal distribution; therefore, the responses were independent and followed a normal distribution. To assess linearity and homoscedasticity or constant variance, scatterplots were constructed showing the standardized residuals versus the standardized predicted values. The standardized residuals scattered randomly about a horizontal line, suggesting constant variance. The scatterplot results suggested a linear pattern. These assumptions for multiple regression appeared to be tenable.

In addition, multicollinearity was also examined. Multicollinearity, high correlations among the predictors, was determined by examining variance inflation factors for the study variables (Stevens, 1996). None of the variance inflation factors exceeded 10. Myers (1990) indicated 10 as a value great enough to cause concern. Multicollinearity, therefore, was not an issue. Data analysis results appear in the following section.

Results

Study respondents included 180 elementary school principals whose level of education varied; 16 had a Masters degree, 111 had received Rank 1 principal certification (approximately 30 hours beyond the Masters degree), 42 had a Specialist degree, and 11 had a Doctorate degree. The respondents included 104 female principals (58%) and 76 male principals (42%), which was very similar to the population of elementary school principals in this state (female principals 52% and male principals 48%). Regarding leadership variables for principals, the mean scores were (a) *Defining the School's Mission* (43.20) from a total of 50, (b) *Managing the Instructional Program* (65.52) from a total of 75, and (c) *Promoting a Positive School Learning Climate* (104.09) from a total of 125. Other descriptive statistics are presented in Table 1.

The multiple regression results appear in Table 2 to answer the research question "What were the significant predictors of student achievement measured by the state assessment?" Standard multiple regression where all the predictor variables are entered into the equation simultaneously was conducted to determine the significance of the equation. Predictor variables entered were: (a) highest level of education obtained by the principal, (b) years of principal experience, (c) years of teaching experience by the principal, (d) years of principal experience at present site, (e) principal gender, (f) principal leadership as determined by the three subscales of the *PIMRS*: *Defining the School's Mission, Managing the Instructional Program, and Promoting a Positive School Learning Climate* and, (*g*) free/reduced lunch population at the school. The criterion variable for multiple regression analysis was student achievement measured by the state assessment. The tables present the standardized regression coefficients (β), levels of significance (*t*), and multiple coefficients of determination (R^2).

The multiple correlation was $R = .53$ with the $R^2 = .28$, indicating that approximately 28% of the variation in the dependent variable, state assessment scores, can be accounted for by the linear combination of independent variables. The adjusted R^2 (.24) was close in degree to R^2 and demonstrated that the variance linked to sampling error was small. Post-hoc statistical power calculations indicated an observed power of .99, which is considered high. Additionally, the obtained effect size, $f^2 = .38$, was large according to Cohen (1988).

The obtained regression equation for principals indicated that one variable, free and reduced lunch ($p < .01$), was a significant predictor of state assessment scores. The beta value for free and reduced lunch was larger ($\beta = -.50$) than any other predictor variables. None of the other variables produced statistical significance.

In summary, the multiple regression procedure indicated that one predictor variable, free and reduced lunch, produced statistical significance regarding student achievement. How-

Table 1
Descriptive Statistics for Principals (N = 180)

Variable	Mean	SD	Range
Background Variables			
Age	48	7.0	29-65
Teaching Experience	14	5.8	2-34
Total Principal Experience	10	5.8	4-25
Principal Experience at Current Site	8	4.9	4-23
Free and Reduced Lunch Population	55	20.6	3-96
Leadership Variables			
Defining the School's Mission	43	4.8	15-50
Managing the Instructional Program	66	6.6	28-75
Promoting a Positive School Learning Climate	104	11.7	47-125
Total *PIMRS* Score	213	20.6	143-250

Table 2
Regression Results: Individual Predictors for State Assessment Scores (N = 180)

Variable	β	R^2	*t*
Gender	.09	.01	1.29
Principal's experience	.00	.00	.04
Teacher's experience	.04	.00	.50
Years in present position	.06	.00	.54
Highest level of education	.01	.00	.08
Free and reduced lunch program	–.50	.25	–7.48**
Defining the School's Mission	.11	.01	1.05
Managing the Instructional Program	–.08	.00	–.77
Promoting a Positive School Learning Climate	–.06	.00	–.59

**$p < .05$.

ever, the remaining variables, highest level of education obtained by the principal, years of principal experience, years of teaching experience by the principal, years of principal experience at present school site, principal gender, and principal leadership as determined by the three subscales of the *PIMRS* did not significantly impact student achievement at the schools participating in this study. A discussion of the implications and conclusions for these results follows in the next section.

Discussion

Not surprisingly, the free and reduced lunch variable in the multiple regression analysis largely accounted for the variance of elementary school state assessment scores. Previously reviewed studies confirm this finding. Bista and Glasman (1998) reported that schools with predominantly low socioeconomic status (SES) were likely to have lower student achievement. Bulach, Malone, and Castleman (1995) found a significant correlation between student achievement and socioeconomic status. Last, the Slovacek, Kunnan, and Kim (2002) study of California schools indicated there was a 2.6 point decline on the state assessment for each percentage point of the student free and reduced lunch population. These previous studies, along with the results of the current one validate the impact free and reduced lunch populations have on student achievement.

In contrast to previous findings from the literature, none of the factors related to principals, highest level of education obtained by the principal, years of principal experience, years of teaching experience by the principal, years of principal experience at present site, principal gender, and principal leadership determined by the *PIMRS* produced statistical significance regarding student achievement on the state assessment. More specifically, (a) a principal's highest level of education did not significantly predict student achievement in contrast with findings from the literature, (Grady & O'Connell, 1993; Gross & Herriott, 1965; Hallinger & Murphy, 1985; Hemphill, Griffiths, & Frederiksen, 1962), (b) years of principal experience did not produce statistical significance unlike several studies (Bista & Glasman, 1998; Hallinger & Murphy, 1985; Young, 1993), (c) principal experience at the present school site did not impact student achievement counter to other research (Bista & Glasman, 1998; Young, 1993), (d) years of principal teaching experience did not predict student achievement and contradicts the literature (Bista & Glasman, 1998; Grady & O'Connell, 1993; Shakeshaft, 1989), (e) principal gender did not predict student achievement distinct from previous research (Brown, Grayson, Brunner, Grogan, & Hackney, 2006; Eagly, Karau, & Johnson, 1992; Kochan, Spencer & Mathews, 2000; Shakeshaft, 1989), and (f) principal leadership skills were not a statistically significant predictor of scores on the state assessment, which contrasted the literature (Hallinger & Heck, 1996; Marzano, Waters, & McNultry, 2005; Spillane & Thompson, 1997; Waters & Grubb, 2004).

With the demands of the No Child Left Behind Act, the answer to what factors impact student achievement continues to merit careful consideration as this legislation demands greater measurement of student achievement and requires that all students make achievement progress. Quite possibly, during this current reform cycle, the complexities of the school organizational structure provide challenges for identifying specific factors that produce increased student achievement. This study, however, expanded the knowledge base about predicting elementary school student achievement using principals and school factors as it confirms the impact of free and reduced lunch on student achievement yet it provides some contradictions with the literature concerning attributes related to principals and their relationships with student achievement.

Perhaps the focus on increasing student achievement should be expanded beyond that of the principal's role. Zmuda, Kuklis and Kline (2004) contend that to improve and transform school structures and meet the high stakes accountability requirements, *leaders* need to "assert the importance of changing minds, not just practices, through the messy processes of dialog, debate, and reflection" (p. vi). Leithwood, Jantzi, and Steinbach (2000) further assert that the decision-making process of the group [principal and teachers] ought to be the central focus for school leaders. Hence, further research that examines instructional leadership

and the collaborative efforts between teachers and principals may help schools understand how this dynamic relates to student achievement. Understanding teacher and principal productivity as it relates to increased student achievement is a worthy research goal and warrants continued interest from educational researchers, policymakers, and practicing school administrators.

Further studies that define instructional leadership appear necessary. Although many characteristics of instructional leadership identified by Hallinger and Murphy (1985, 2001) are present in the widely adopted Interstate School Leadership Licensure Consortium standards used to design university school administrative programs and for state administrative licensure across the country, there is no single accepted description or definition of the principal's role as an instructional leader (Marzano, Waters, & McNulty, 2005). Marzano, Waters, and McNulty (2005) point out that "Despite its popularity, the concept [instructional leadership] is not well defined" (p. 16). Having a clear definition of instructional leadership and school stakeholders who are involved with instructional leadership is a worthy research goal and warrants continued interest from educational researchers, policymakers, and practicing school administrators.

References

Babbie, E. R. (1990). *Survey research methods*. Belmont, CA: Wadsworth Publishing.

Bista, M., & Glasman, N. (1998). Principals' approaches to leadership, their antecedents and student outcomes. *Journal of School Leadership*, *8*, 109-136.

Bulach, C. R., Malone, B., & Castleman, C. (1995). An investigation of variables related to student achievement. *Mid-Western Educational Researcher*, *8*(2), 23-29.

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

Dillman, D. A. (2000). *Mail and internet surveys*: *The tailored design method*. New York: John Wiley & Sons.

Eagly, A. H., Karau, S. J., & Johnson, B. T. (1992). Gender and leadership style among school principals: A meta-analysis. *Education Administration Quarterly*, *28*, 76-102.

Fullan, M. (2002). The change leader. *Educational Leadership*, *59*(8), 16-20.

Grady, M. L., & O'Connell, P. A. (1993). Women in K-12 educational administration: A synthesis of dissertation research. *Journal of School Leadership*, *3*, 449-460.

Gross, N., & Herriott, R.E. (1965). *Staff leadership in public schools: A sociological inquiry.* New York: John Wiley & Sons.

Hallinger, P. (1985). *Resource manual: The principal instructional management rating scale.* Nashville, TN: Vanderbilt International Institute for Principals.

Hallinger, P., & Murphy, J. (1985). Assessing the instructional management behavior of principals. *The Elementary School Journal*, *86*(2*)*, 217-247.

Hallinger, P., & Heck, R. (1996). Reassessing the principal's role in school effectiveness: A review of the empirical research, 1980-1995. *Educational Administration Quarterly*, *32*, 5-44.

Hallinger, P., & Taraseina, P. (2001, April). *The principal's role as instructional leader: A review of studies using the PIMRS.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Hemphill, J. K., Griffiths, D. E., & Frederiksen, N. (1962). *Administrative performance and personality.* New York: Teachers College Bureau of Publications.

Kochan, F. K., Spencer, W. A., & Mathews, J. G. (2000). Gender-based perceptions of the challenges, changes, and essential skills of the principalship. *Journal of School Leadership*, *10*, 290-310.

Latham, G., & Wesley, K. (1981). *Increasing productivity through performance appraisal.* Menlo Park, CA: Addison Wesley.

Leithwood, K., Begley, P., & Cousins, P. (1990). The nature, causes and consequences of principals' practices: An agenda for future research, *Journal of Educational Administration*, *28*(4), 5-31.

Leithwood, K., Jantzi, D., & Steinbach R. (2000). Changing leadership: A menu of possibilities. In K. Leithwood, D. Jantzi, & R. Steinbach (Eds.), *Changing leadership for changing times* (pp. 3-20), Philadelphia: Open University Press.

Marzano, R. J., Waters, T., & McNultry, B. A. (2005). *School leadership that works: From research to results.* Aurora, CA: Mid-continent Research for Education and Learning.

Murphy, J. (1992). *The landscape of leadership preparation: Reframing the education of school administrators.* Newbury Park, CA: Sage Publications.

Myers, R. (1990). Classical and modern regression with applications (2nd ed.). Boston: Duxbury Press.

Nunnally, J. (1967). *Psychometric theory*. New York: McGraw-Hill.

Shakeshaft, C. (1989). The gender gap in research in educational administration. *Educational Administration Quarterly*, *25*, 324-337.

Shakeshaft, C., Brown, G., Grayson, D., Brunner, C., Grogan, M., & Hackney, C. (2006, November). *Gender and leadership: A conversation among researchers.* Paper published at the annual meeting of the University Council for Educational Administration, Nashville, TN.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*, 417-453.

Slovacek, S., Kunnan, A., & Kim, H. (2002, March). California charter schools serving low-SES students: An analysis of the academic performance index. Paper presented at the California Network of Educational Charters Conference, Los Angeles.

Spillane, J. P., & Thompson, C. C. (1997). Reconstructing conceptions of local capacity: The local education agency's capacity for ambitious instructional reform. *Educational Evaluation and Policy Analysis*, *19*, 185-203.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Waters, T. & Grubb, S. (2004). *Leading schools: Distinguishing the important from the essential.* Aurora, CA: Mid-continent Research for Education and Learning.

Young, J. H. (1993). Principals and collaborative curriculum development: Does gender make a difference? *Alberta Journal of Educational Research*, *4*, 433-448.

Zmuda, A., Kuklis, R. & Kline, E. (2004). *Transforming schools: Creating a culture of continuous improvement.* Virginia: ASCD.

Sharon Gieselmann is an Assistant Professor of Education at the University of Evansville. Her research interests include equity and social justice issues relating to student achievement. E-mail: sg78@evansville.edu.

Call for Manuscripts

The ***Mid-Western Educational Researcher*** is a scholarly journal that publishes research-based articles addressing a full range of educational issues. The journal also publishes literature reviews, theoretical and methodological discussions that make an original contribution to the research literature, and feature columns. There are four issues of the journal published annually.

Manuscripts are accepted from faculty, students, and professionals working in educational or non-educational settings. Membership in the MWERA is not required in order to submit a manuscript for review. The editors encourage the submission of revised papers that have been presented at the annual meetings of the MWERA, AERA, and other professional organizations.

In addition, the editors encourage quality graduate student submissions through a devoted section, "Graduate Student Research," for papers authored by graduate students (as either sole or first author). This does not preclude manuscripts authored by graduate students from appearing in the main section of the journal. The submission, review, and publication of manuscripts in this section conform to the descriptions and standards of the journal outlined below. Manuscripts should be submitted electronically with *MWER Graduate Student Manuscript* as the subject line. It is essential to identify yourself as a graduate student to be considered for the Graduate Student Section. Verification of graduate student status will be required if the manuscript is accepted for publication in *MWER*.

All manuscripts must be submitted electronically for review to Dr. Julia Matuga, Submissions Co-Editor, at mer@bgsu.edu as an email attachment. Indicate in the subject line that this is a MWER manuscript. Manuscripts should be formatted as an MS Word document using 12 point Times New Roman font. Manuscripts should conform to the style and format described in the *Publication Manual of the American Psychological Association, 5th edition*. All manuscripts should be typed, double-spaced, with 1½ inch margins on all sides, and include page numbers. An abstract of fewer than 100 words should accompany the manuscript. The author's name, contact information, and affiliation should appear on the title page only. Submissions typically are fewer than 20 pages in length, including references, title page, and abstract.

All manuscripts will be acknowledged electronically upon receipt. Please note that authors are responsible to submit manuscripts that are free of grammatical and mechanical errors. Manuscripts will be screened initially for format and fit for the journal by the editors. Appropriate manuscripts will be submitted to blind review. The editors reserve the right to make minor modifications in order to produce a more concise and clear article. Contributors acknowledge by virtue of their submission to the journal that they will consent to have their work available internationally through the EBSCO portal, as per agreement with the MWERA.

Questions regarding the journal should be directed to the Submissions Co-Editor:

Dr. Julia M. Matuga
College of Education & Human Development
Bowling Green State University
444 Education Building
Bowling Green, OH 43403
(419) 372-7317
mer@bgsu.edu

The Mid-Western Educational Research Association's

32nd Annual Meeting
October 14–17, 2009

Sheraton Westport Chalet Hotel
St. Louis, Missouri

The Synthesis of Educational Research and Practice

The 2009 Annual Meeting of the ***Mid-Western Educational Research Association*** will be held in St. Louis with an exciting program of invited speakers, focused workshops, and peer-reviewed papers presented in a variety of session formats. We will kick off the program with our traditional Fireside Chat with **Dr. Ralph Martin**, Professor Emeritus of Teacher Education at Ohio University, who will also be giving our keynote address on Thursday. Our Friday luncheon speaker is **Dr. Ronald Rochon**, Dean of the School of Education and Associate Vice President for Teacher Education at Buffalo State College. We are very fortunate to have two such outstanding individuals giving our invited addresses. Teachers, administrators, and other school personnel are invited to come and share their work and experiences within this theme, as well as other topics, at the 2009 MWERA conference. Educational researchers across North America will once again return to MWERA to renew acquaintances, make new contacts, and engage in exciting conversation in a collegial atmosphere.

Come and be a part of MWERA 2009 in St. Louis!

Conference and hotel registration information and forms are available at the MWERA web site:

http://www.mwera.org

Thursday Morning Keynote Address

Featured Speaker

Dr. Ralph Martin

Dr. Ralph Martin (Ph.D., University of Toledo) is Professor Emeritus of Teacher Education at Ohio University, Athens. During his 35 years as an educator, Dr. Martin has taught science lessons and led projects spanning K-16. His research and scholarship has produced more than 15 books and manuals, including the popular textbook *Teaching Science for All Children: An Inquiry Approach* (in its 5th edition). He has obtained millions in grant funding for teacher professional development in science and math. Dr. Martin has received numerous Ohio and national awards in recognition of his teaching and leadership and has coordinated, directed, and chaired many university programs and faculty units. Presently he serves as Chair of the Science Review Board for the Ohio Resource Center for Mathematics, Science and Reading (www.ohiorc.org) and is Co-Director of the South Eastern Ohio Center of Excellence for Mathematics and Science (www.seocems.org), which conducts research, provides professional development programs, and supports educational improvement efforts focused on Appalachian Ohio.

Friday Luncheon Keynote Address

Featured Speaker

Dr. Ronald Rochon

Dr. Ronald Rochon (Ph.D., University of Illinois at Urbana-Champaign), currently serves as the Dean of the School of Education and Associate Vice President for Teacher Education at Buffalo State College in Buffalo, NY. Dr. Rochon's areas of interest include the recruitment, retention, and successful matriculation of culturally, linguistically, and racially diverse students; strengthening the relationship between communities and universities; and assisting students in examining and deepening their understanding of the ways in which socio-cultural factors influence American educational thought, theory, and practice. Dr. Rochon's research and teaching address the historical and contemporary perspectives surrounding the politics of race and culture within American society. He is committed to community renewal and collaboration. Dr. Rochon works extensively with stakeholders in PK-12 settings and throughout the community to develop systemic community building and empowering relationships. Throughout his tenure, Dr. Rochon has secured millions in private, state, and federal grant funding to establish educational programming and workshops, scholarships for underserved populations, and international exchange programs.

Sheraton Westport Chalet Hotel St. Louis – Lakeside Chalet

191 Westport Plaza, St. Louis, Missouri 63146 (Phone: 314-878-1500)
http://www.sheratonwestport.com/westport.html

Nestled in St. Louis' impressive West Port Plaza area, the Lakeside Chalet hotel is situated just five miles from Lambert International Airport, with complimentary airport transportation. Westport Plaza features 18 restaurants, 2 nightclubs, a comedy club and several specialty shops. The Sheraton Westport Lakeside Chalet is centrally located for all your plans to visit St. Louis attractions and is just minutes from historic St. Charles, the Saint Louis Zoo, Art Museum, and the History Museum. The world famous Gateway Arch, Harrah's St. Louis Riverport Casino, Ameristar Casino, Westfield Shopping Town, St. Louis Mills Mall, and the St. Louis Galleria Shopping Center are all nearby.

Stay fit with state-of-the-art fitness centers and swimming pools at the Sheratons at Westport. Stay connected with free Wi-Fi and free PC workstations. Guestroom amenities include:

- Sheraton Sweet Sleeper® Beds, now available in all guestrooms
- Wireless High Speed Internet Access in all guestrooms
- 27" Television with Cable including ESPN, CNN and HBO, and In-room movies
- Coffee Maker, Iron and full size Ironing Board, Hair Dryer and Bath Amenities
- Complimentary USA Today Newspaper delivered Mon-Fri
- Oversized Desk and Executive Leather Chairs

REGISTRATION:
Use the adjacent form, or go to:
http://www.mwera.org

The Sheraton Westport Hotel Reservation Form
Mid-Western Educational Research Association Meeting
October 14–17, 2009

Your Name: _____
(First Name) (Middle Initial) (Last Name)
Affiliation: _____
Mailing Address: _____
Day Telephone: (_____)_____ E-mail: _____

Accommodations Requested
Arrival Date: _____/_____/ 2009 Departure Date: _____ / _____ / 2009
Bed Type: ____ Single ____ Double (2 Doubles)

- ❑ $135—Standard single conference rate
- ❑ $135—Double Rate (two double beds)

Name(s) of Roommate(s) (if any): _____

Special Needs: _____

Confirming your reservation requires a first night's deposit or a credit card.

Method of Payment
- ❑ Check or Money Order
- ❑ Credit Card (*Circle to indicate card*):
  MasterCard Visa American Express
  Discover Diners Club

Credit Card Number: _____
Name on Credit Card: _____
Expiration Date: _____
Signature: _____

You must cancel this reservation 72 hours prior to your expected date of arrival and receive a cancellation number to avoid billing on your credit card for the first night's room and tax or the loss of your deposit. The above rates do not include state and local taxes. Automobile parking (valet or self-parking) is available at the hotel for an additional $20 per day (plus taxes) for registered hotel guests. Check in time is 3:00 pm; check out time is 1:00 pm. On site luggage storage is available for early arrival and late check out.

The above group rates are only guaranteed UNTIL SEPTEMBER 13, 2009.

Phone or send completed form and deposit
by mail or fax to:
Sheraton Westport Chalet Hotel St. Louis
191 Westport Plaza
St. Louis, Missouri 63146
Phone: 314-878-1500
Reservations only: (800) 822-3535
http://www.sheratonwestport.com/westport.html
Be sure to mention "MWERA" when making your reservation!

Please support the conference by reserving your room at the Sheraton. MWERA reserves a block of discounted rooms for attendees. If these rooms are not booked, the conference must pay a sizeable penalty.

clip and send ✂

The Teacher Work Sample: Candidate and Mentor Perceptions

David Bell
Maureen Spelman
Holly Mackley
Liang Zhao
Saint Xavier University, Chicago

Abstract

Elementary education faculty at this university embedded a Teacher Work Sample (TWS) performance-based assessment into the student teaching experience to assess candidates' ability to impact student learning as required by recent accreditation reforms. The authors conducted an internal evaluation of the TWS for the purposes of strengthening not only the capstone experience, but also all of the courses and field experiences that precede student teaching. This study examines quantitative and qualitative data gathered via surveys and follow-up semi-structured interviews with candidates and mentors. The results of the study indicate that while both candidates and mentors perceive the TWS as a positive tool, mentors had a significantly more positive view than candidates in several crucial areas.

Prior to the mid 1990s, the emphasis in teacher preparation programs was on the process of becoming a teacher. Schools of education were concerned with how teacher-candidates (called "candidates" throughout) learned to teach and how their beliefs and attitudes evolved throughout their educational coursework. Instructors were concerned with providing pedagogical knowledge in contexts that supported candidate learning. More recently, however, teacher preparation programs have entered into a new paradigm. Universities are now faced with critical examination from external auditors assessing the qualifications of faculty members, the alignment of courses and field work with national and state standards, and candidate impact on student learning.

Teacher preparation has moved from an input approach focused on preservice teaching credentials toward an outcomes approach measuring candidates' classroom performance; this reframing of accountability has placed teacher preparation programs in the position of demonstrating that their candidates can, indeed, make a positive impact on P-12 student learning (Rothman, 2008). Such demands have even appeared in federal legislation as a proposal that professional teacher preparation programs be evaluated, at least in part, on the academic achievement of the P-12 students of their graduates (Hamel & Merz, 2005). In response to this shift toward increased accountability, schools of education have begun to pilot ways to assess candidates' impact on P-12 student achievement (Rothman, 2008). The path from policy into practice now starts with schools of education: administrators redesign curricula and align courses with reform policies and standards; instructors strive to build the knowledge, skills, and dispositions of their candidates in the belief that those candidates will, in turn, impact student learning in the P-12 schools (Spelman, 2006).

Theoretical Framework

Educational Reform and Teacher Quality

Beginning with *A Nation at Risk,* educational reforms in America experienced a shift in focus that has resulted in two decades of national and state mandates aimed at improving teaching and learning in P-12 schools. This increased federal focus on teacher quality has occurred, at least in part, because more and more research indicates that teacher expertise and the quality of instruction are highly significant factors in determining student achievement (Darling-Hammond & Ball, 2002; Greenwald, Hedges, & Laine, 1996; Ingersoll, 1996; Zumwalt & Craig, 2005). In 1996, the National Commission on Teaching and America's Future (NCTAF) was given the charge to study changes needed to ensure that every child in the United States would have access to highly qualified teachers. NCTAF (1996) recommended that schools of education work with states to redesign teacher education programs so that all candidates have access to high-quality learning opportunities.

The updated NCTAF (2003) report recommended that individual states require all teacher preparation programs to meet rigorous accreditation standards, establish institution-wide and program-wide leadership responsibility for the quality of teacher preparation, and if necessary, close those programs that are unable to produce high-quality teachers. In a move designed to raise the quality of the next generation of teachers, the Commission challenged institutions of higher education to collect and use data on P-12 student achievement, teacher licensure, and teacher retention to improve teacher preparation programs. In essence, university teacher education programs are being asked to ensure that the new professional has the knowledge necessary for effective

classroom practice, is prepared for stringent initial licensure tests, and can demonstrate learning gains for all P-12 students (Cochran-Smith, 2005; Wise & Leibbrand,1996). This increased attention to student learning as a measure of teacher effectiveness has moved the preparation of teaching professionals into the spotlight (Girod & Girod, 2006).

Effective Teacher Preparation

To support the development of exemplary teacher preparation programs, NCTAF worked with its state partners to build upon the research and standards developed by the Interstate New Teacher Assessment and Support Consortium (INTASC) and the National Board for Professional Teaching Standards (NBPTS) to develop a consensus about what a highly qualified beginning teacher should know and be able to do to help students learn. In this consensual vision, teacher education programs would be redesigned to ensure that candidates acquire a thorough knowledge base in a balanced program that places equal emphasis on content knowledge, pedagogical knowledge, and pedagogical content knowledge (Shulman, 1987).

Effective teachers are frequently described as those who are willing to reflect on practice to improve teaching (Cuban, 1990). To develop these personally motivated teachers, teacher preparation programs need to provide multiple opportunities for teacher candidates to engage in performance-based formative assessments. The state of Connecticut has led the way toward performance assessments for those candidates seeking initial licensure. More recently, thirty-one teacher preparation programs in California have switched to a performance-based assessment, The Performance Assessment for California Teachers (PACT), as a means of measuring the classroom performance of candidates (Rothman, 2008). The Renaissance Partnership Teacher Work Sample (TWS) has also emerged as a performance-based assessment which synthesizes professional education coursework into a comprehensive unit that examines candidate impact on P-12 student learning (Delvin-Scherer, Daly, Burroughs, & McCartan, 2007).

The Teacher Work Sample Methodology

Educators at Western Oregon University designed an approach in which candidates are explicitly taught a model for teaching and learning that involves a design for effective planning, instruction, assessment, and reflection with an emphasis on assembling and analyzing data on P-12 student learning. This model evolved into seven processes (i. e., contextual factors, learning goals, assessment plan, instructional design, instructional decision making, and reflection on teaching and learning) commonly known as the TWS Methodology (Girod, 2002; Denner, Norman, Salzman, & Pankratz, 2003). This performance-based assessment was adapted by the Renaissance Group, a consortium of eleven colleges and universities, in their joint initiative to improve teacher quality through connecting teacher performance to student learning (Pankratz, 1999).

The TWS is a vehicle that guides candidates' thinking about the processes of teaching in ways that are tightly linked to P-12 student learning. When implemented as a means of gaining teaching experience in this manner and demonstrating effectiveness in doing so, a performance-based TWS can be considered both a vehicle to guide instruction as well as an approach to measurement (Girod, 2002). In fact, when used as an instructional framework, TWS Methodology scaffolds candidates as they question and reflect upon their teaching decisions. As an approach to measurement and accountability, the TWS allows candidates to examine student learning on specific outcomes that have been the focus of the instruction. It can also allow candidates to place student progress in a contextually grounded portrayal that supports analysis of student learning and candidate teaching in an authentic setting. This performance-based assessment examines a candidate's work and a work of P-12 students, thus providing a way of meaningfully connecting the two samples (Schalock & Myton, 2002).

Context

Teacher preparation has become the joint responsibility of numerous stakeholders, and yet current research is dominated by the voices and perceptions of university faculty and administrators. There is a need to examine the voices, perceptions, and questions raised by candidates and their mentors in the field (Cochran-Smith & Zeichner, 2005). The researchers of this study designed such an investigation to examine the perceptions of candidates and mentors regarding the benefits and challenges of implementing the TWS during the capstone student teaching experience. This midwestern university is a private, faith-based, liberal arts institution located in a major metropolitan setting. Founded in the 1860s, the university serves approximately 5,700 students at the main campus and off-site locations. Initial teacher certification programs are offered at both the graduate and undergraduate levels.

Historically, this university's elementary education program required candidates to plan and implement a unit of study during their 16-week student teaching experience. Candidates were also responsible for conducting an action-research study in their clinical practice classroom. However, faced with the new conditions for program approval, the elementary education faculty began to search for a tool that would evaluate candidates by focusing on student learning as well as candidate performance. This search led a team of the researchers to a Renaissance Group workshop focused on the performance-based Renaissance TWS. An initial examination of the TWS Methodology revealed several common features already in place in the traditional capstone assignments at this university.

As a result of the experience at the workshop, the elementary education program faculty proceeded to adapt a TWS performance-based assessment to be used as a tool for professional preparation, performance accountability, and program improvement. This customized version of the TWS

was piloted in the capstone student teaching experience during the fall 2005 academic term. The TWS became one of the core assessments used to demonstrate candidates' knowledge, skills, dispositions, and impact on student learning in the accreditation report to the Association for Childhood Education International (ACEI). Currently, the TWS serves as one piece of evidence of candidates' ability to impact their students' achievement prior to exiting the program. This study is part of a larger, ongoing internal evaluation of the TWS aimed at strengthening not only the capstone experience, but also all of the courses and field experiences that precede student teaching.

Methodology

A review of the literature reveals that the voices and perceptions of university faculty and administrators have dominated TWS studies. Few studies have been conducted with a focus on the perceptions of the largest group of stakeholders: candidates and their mentors. This study has been designed to address that gap in the TWS literature. Participants in this study included candidates enrolled in an elementary graduate or undergraduate initial certification program during the fall or spring academic terms of 2006. Specifically, the study focused on candidates enrolled in the program's capstone course (i. e., student teaching). The student teaching course is the last in a series of professional education courses that must be completed prior to certification. Each candidate was assigned at least one site-based teacher in a K-8 classroom mentor who supported and facilitated the student teaching experience and worked daily with candidates on critical design components as well as the implementation of the TWS. Thus, as significant stakeholders, it was important to also gather data regarding mentors' perceptions on the impact of the TWS on raising candidates' performance.

A blend of quantitative and qualitative measures was used to triangulate the data and gain a better understanding of participants' perceptions of the TWS on raising the quality of candidate performance. Specifically, the researchers sought to answer the following questions: (a) what are the perceptions of candidates regarding the impact of the TWS on raising the quality of their instructional effectiveness? (b) what are the perceptions of mentors regarding the impact of the TWS on raising the quality of candidate performance? and (c) is there a significant difference between the perceptions of candidates and mentors regarding the impact of the TWS on raising the quality of candidate performance? To answer these questions the researchers collected data using a Likert scale survey instrument, open-ended questionnaires, and semi-structured interviews during the spring and fall 2006 academic terms.

Data Collection

During the spring and fall of 2006, a 5-point, 17-item Likert scale survey instrument was used to collect data regarding candidate and mentor perceptions of the TWS. Participants responded by assigning an answer of 1 = strongly disagree (SD), 2 = disagree (D), 3 = neutral (N), 4 = agree (A), or 5 = strongly agree (SA). Survey instruments were distributed to candidates at the final class session. Although the completion of the survey was voluntary, the response rate was 100% with a final sample size of 107. To solicit the perceptions of mentors, researchers mailed 112 survey instruments. The response rate was 36% with a final sample size of 40. To determine the internal consistency of the data received, a Chronbach reliability analysis was conducted on candidate and mentor survey responses resulting in an alpha level of .98 and .96, respectively, revealing a high rate of variability. Items on the survey instrument included questions related to the impact of the TWS assignment on candidates' ability to use the seven TWS processes, the development of professional competencies, and whether or not the TWS reflects mentors' current practice at the clinical supervision site.

The survey instrument also included open-ended response items that invited participants to share any concerns or recommendations regarding the TWS and offer additional feedback regarding their perceptions of the TWS. A total of 63 candidates and 16 mentors responded to open-ended items on the survey instrument. The final item on the survey instrument was a short paragraph inviting respondents to participate in a follow-up interview. Fifteen volunteers, 11 out of 107 candidates and 4 out of 40 mentors, were willing to share their perceptions in follow-up semi-structured interviews (Patton, 2002; Rossman & Rallis, 2003). The face to face or telephone interviews lasted approximately forty-five minutes and were guided by parallel protocols built upon questions regarding candidate effectiveness in implementing the seven TWS processes as well as a number of supporting probes designed to encourage participants to expand on their responses (Rossman & Rallis, 2003). Participants were also encouraged to share their perceptions regarding the overall strengths and challenges of implementing a TWS during the student teaching experience.

Data Analysis

A significant amount of both qualitative and quantitative data were gathered through survey instruments, open-ended questionnaires, and semi-structured interviews. The tasks of managing and analyzing that data required the use of multiple methods of data analysis. Initially, a frequency distribution was employed to examine perceptions of the survey participants regarding the impact of the TWS on raising candidate performance. Because the responses on the Likert scale survey were considered ordinal data and the independent variables (candidates and mentors) were considered categorical data, a nonparametric Mann-Whitney test was selected. A Mann-Whitney test was chosen over a chi-square as an appropriate statistical test for two reasons: (a) the sample size of mentors was small (n = 40), and (b) the researchers did not make any assumptions regarding the distribution of the population. The use of a Mann-Whitney test allowed researchers to evaluate whether the median on

a test variable differed significantly between the two groups of participants.

To analyze the open-ended response items on the survey instrument and the semi-structured interviews, the researchers reviewed the responses of mentors using the same 5-point Likert scale rankings found in the survey. This allowed a comparison between the quantitative and qualitative data from the survey instrument. Candidates' open-ended response items and the semi-structured interviews, however, were analyzed using two discrete methods of analysis. Initially, candidate responses were scored using the same Likert scale rankings used to analyze the responses of mentors. To gather additional data regarding teacher candidates' progress towards reflective practice, a second analysis was completed. This second analysis involved the use of a rubric aligned with Carol Rodgers' (2002a) reflective cycle based on Dewey's criteria (1910/1933). In this second analysis the researchers reviewed the statements of candidates and evaluated those responses against a rubric designed to measure their growth toward the ultimate goal of reflective practice. The holistic rubric values and criteria for placement in the reflective cycle were as follows:

Level 0. superficial, comments not related to the professional growth experience

Level 1. presence in the experience, learning to see as a state of mindfulness, full awareness and concentration, learning centered

Level 2. description of the experience, learning to describe and differentiate, slowing down to look and see the variety of nuances present

Level 3. analysis of the experience, learning to think from multiple perspectives and form multiple explanations, reorganizing and reconstructing the experience

Level 4. experimentation, learning to take intelligent action, testing ideas, taking risks

Results

Survey-Instrument Questions

Candidate and mentor responses on each item of the survey were averaged and compared to understand their overall perceptions of the impact of the TWS on raising candidates' performance. Statistics reveal that mentors perceive the impact more positively than candidates on every item on the survey. Specifically, mentors' mean scores ranged from 3.58 – 4.15 revealing an overall favorable perception of the TWS. Candidates' mean scores ranged from 2.63 (question 16) – 3.89 (question 5), revealing mixed perceptions on the impact of the TWS. Table 1 provides complete details of the descriptive statistics.

To understand the distribution of mentor and candidate responses to the survey items, a frequency table was used to summarize and organize data. Frequency distribution test results in Table 2 indicate that over 50% of mentors responded positively regarding the impact of the TWS on raising the quality of candidates' performance. With respect to candidate perceptions (when combining *strongly agree* and *agree*), Table 2 indicates over 50% of candidates responded positively (when combining *strongly agree* and *agree)* to 13 of the 17 items regarding the impact of the TWS on the quality of candidates' performance. Conversely, less than 50% of candidates indicated negative responses (when combining *agree* and strongly agree) to the remaining four items (questions 13, 15, 16, and 17) on the survey. Table 2 summarizes the responses to the survey.

Table 1
Descriptive Statistics of Candidate with Respect to Each Question on the Likert Scale Survey

Question Focus	IV	M	SD
Understand information to…			
1. plan instruction	C	3.25	1.21
	M	3.87	1.04
2. guide assessment plan	C	3.00	1.12
	M	3.80	1.04
Create challenging…			
3. learning goals	C	3.90	1.27
	M	4.10	1.13
Use learning objectives to develop…			
4. pre-assessment plan	C	3.70	1.17
	M	3.90	1.10
5. formative assessment	C	3.89	1.14
	M	3.97	1.05
6. post-assessment	C	3.75	1.16
	M	4.05	1.01
Design instruction consistent with…			
7. student information	C	3.75	1.19
	M	4.05	1.02
8. objectives	C	3.51	1.15
	M	4.05	1.01
9. pre-assessment plan	C	3.68	1.22
	M	4.19	1.12
Conduct formative assessment to…			
10. modify instruction	C	3.46	1.23
	M	3.92	1.07
Ability to reflect…			
11. on student learning	C	3.70	1.26
	M	4.00	.99
12. after instructional delivery	C	3.62	1.34
	M	4.15	.89
Overall the TWS…			
13. demonstrates effective teaching	C	3.09	1.34
	M	3.90	.89
14. structure supports student learning	C	3.19	1.34
	M	3.80	1.15
15. supported my growth as a professional educator	C	3.07	1.24
	M	3.75	1.26
16. raised the quality of candidate performance	C	2.63	1.27
	M	3.58	1.30
17. reflected current clinical practice site	C	3.03	1.37
	M	3.83	1.08

Note: IV = independent variable; C = candidate; M = mentor

Tables 3 displays the results of a Mann-Whitney test conducted to determine if there was a significant difference in the perceptions of candidates and mentors regarding the impact of the TWS on raising the quality of candidate performance. Table 3 reveals that mentors perceived the TWS as significantly more positive than candidates did in the areas of instructional planning (question 1), with a mean place of 89.54 and 68.19, respectively, and creating learning goals (question 2), with a mean place of 95.31 and 66.03, respectively. Table 3 provides detail of the data results. Mentors perceived the TWS as significantly more positive than candidates did in designing appropriate instruction consistent with learning objectives (question 8), with a mean place of 84.06 and 68.20, respectively. Furthermore, mentors perceived the TWS as significantly more positive than candidates did in designing appropriate instruction consistent with a pre-assessment plan (question 9), with a mean place of 84.21 and 68.73, respectively.

Table 3 also reveals that mentors perceived the TWS as significantly more positive than candidates did in the ability to conduct formative assessments to modify instruction. Mentors perceived the TWS as significantly more positive than candidates did in the ability to reflect on instructional delivery (question 12), with a mean place of 84.85 and 69.94, respectively. Furthermore, mentors perceived the impact of the TWS as significantly more positive than candidates did in the areas of effective teaching (question 13), supporting

Table 2
Frequency Distribution of Candidate (C) and Mentor (M) Response to the Items on the Survey

Question Focus	IV	*n*	SD	D	N	A	SA
Use student information to…							
1. plan instruction	C	107	10.0%	19.6%	16.8%	41.1%	12.1%
	M	40	0	15.0%	15.0%	35.5%	32.5%
2. guide assessment plan	C	105	6.5%	13.1%	5.6%	31.8%	41.1%
	M	40	2.5%	12.5%	12.5%	47.5%	25.0%
Create challenging…							
3. learning goals	C	105	6.7%	13.3%	5.7%	32.4%	41.9%
	M	40	5.0%	5.0%	12.5%	30.0%	47.5%
Use learning objectives to develop…							
4. pre-assessment plan	C	105	3.8%	15.2%	19.0%	31.4%	30.5%
	M	40	5.0%	7.5%	12.5%	42.5%	32.5%
5. formative assessment	C	63	3.2%	11.1%	14.3%	36.5%	34.9%
	M	40	5.0%	5.0%	10.0%	47.5%	32.5%
6. post-assessment	C	103	6.8%	7.8%	16.5%	38.8%	30.1%
	M	41	5.0%	2.5%	12.5%	47.5%	32.5%
Design instruction consistent with…							
7. student information	C	100	7.0%	10.0%	14.0%	39.0%	30.0%
	M	39	2.6%	7.7%	10.3%	41.0%	38.5%
8. objectives	C	105	8.6%	11.4%	9.5%	44.8%	25.7%
	M	39	5.1%	5.1%	10.3%	33.3%	46.2%
9. pre-assessment plan	C	105	8.6%	19.0%	8.6%	45.7%	18.1%
	M	40	2.5%	10.0%	15.0%	37.5%	35.0%
Conduct formative assessment to…							
10. modify instruction	C	105	6.7%	16.2%	13.3%	46.7%	17.1%
	M	40	2.5%	7.5%	7.5%	42.5%	40.0%
Ability to reflect…							
11. on student learning	C	104	10.6%	9.6%	5.8%	47.1%	26.9%
	M	40	0%	10.0%	17.5%	35.0%	37.5%
12. after instructional delivery	C	107	12.1%	11.2%	8.4%	39.3%	29.0%
	M	40	0%	7.5%	10.0%	42.5%	40.0%
Overall the TWS…							
13. demonstrates effective teaching	C	107	15.0%	23.4%	15.0%	30.8%	15.9%
	M	40	2.5%	12.5%	17.5%	27.5%	40.0%
14. structure supports student learning	C	107	11.2%	22.4%	15.0%	39.3%	12.1%
	M	40	5.0%	15.0%	15.0%	25.0%	40.0%
15. supported my growth as a professional educator	C	107	16.8%	14.0%	26.2%	30.8%	12.1%
	M	40	7.5%	15.0%	7.5%	35.0%	35.0%
16. raised the quality of candidate performance	C	107	19.6%	28.0%	28.0%	18.7%	5.6%
	M	40	7.5%	15.0%	22.5%	22.5%	32.5%
17. reflected current clinical practice site	C	101	17.8%	20.6%	20.6%	23.4%	17.8%
	M	40	2.5%	12.5%	15.0%	40.0%	30.0%

Note: *n* = sample;
SD = strongly disagree; D = disagree; N = neutral; A = agree; SA = strongly agree

student learning (question 14), becoming a professional (question 15), and developing candidate teaching performance (question 16). Finally, mentors perceived the TWS as significantly more representative of the clinical practice site than did candidates, with a mean place of 91.74 and 67.37, respectively.

Open-Ended Response Items

Sixteen mentors responded to the open-ended survey items. Eleven of the 16 mentor responses to the first open-ended item asking for concerns and recommendations fell into the negative range. Mentors generally focused on the problems and barriers encountered by the candidates as they designed and implemented the TWS: "Even though the TWS was helpful it took a lot of time and effort. (Perhaps) it should be the only (student teaching) requirement."

However, 12 of the 16 mentor responses to the second open-ended item asking for specific feedback on the seven TWS processes fell largely in the positive range. Mentors shifted their focus here to the importance and benefits of the TWS: "TWS helps student teachers understand more about their workplace." Finally, the third opportunity for additional feedback resulted in only positive statements related to candidate performance and the overall mentoring experience.

Table 3
Statistical test to determine the between-group difference

Question Focus	IV	Mean Rank	Sum of Ranks
Use student information to...			
1. plan instruction	C	68.19	7296.50
	M	*89.54	3581.50
2. guide assessment plan	C	66.03	7065.50
	M	*95.31	3812.50
Create challenging ...			
3. learning goals	C	71.48	7505.00
	M	78.19	3080.00
Use learning objectives to develop ...			
4. pre-assessment plan	C	71.02	7457.50
	M	78.19	3127.50
5. formative assessment			
	C	51.31	3232.50
	M	53.09	2123.50
6. post-assessment	C	70.05	7215.00
	M	77.03	3081.00
Design instruction consistent with ...			
7. student information	C	67.32	6731.50
	M	76.88	2998.50
8. learning objectives	C	68.20	7161.50
	M	*84.06	3278.50
9. pre-assessment plan	C	68.73	7216.50
	M	*84.21	3368.50
Conduct formative assessment to ...			
10. modify instruction	C	66.89	7023.50
	M	*89.04	3561.50
Ability to reflect ...			
11. on student learning	C	70.36	7317.00
	M	78.08	3123.00
12. after instructional delivery	C	69.94	7484.00
	M	*84.85	3394.00
Overall the TWS ...			
13. demonstrates effective teaching	C	67.10	7180.00
	M	*92.45	3698.00
14. structure supports student learning	C	68.22	7300.00
	M	*89.45	3578.00
15. supported candidate growth as a professional educator	C	67.71	7245.00
	M	*90.83	3633.00
16. raised the quality of candidate performance	C	65.85	7045.50
	M	*95.81	3832.50
17. reflected mentor's current practice at clinical site	C	67.37	7208.50
	M	*91.74	3669.50

Note: IV = independent variable; C = candidate; M = mentor

Mentors frequently commented that the TWS was helpful but indicated a need for further clarity regarding the assignment requirements and their particular supporting roles. "I would have liked a letter of introduction (explaining) my role." and "It was important to provide clarity with regard to the TWS . . . should it be cross curricular, problem-based, etc.?" A number of mentors suggested that TWS components might be better placed prior to the student teaching capstone experience, while others felt it was beneficial during student teaching. "(The TWS might be) more effective if designed outside the student teaching experience . . .very time consuming." and "While the completed project was excellent, the amount of time and energy that was spent putting it together could have been better spent."

Twenty-nine of the 63 candidate responses to the first open-ended response item were also negative. Comments made by candidates tended to focus on the amount of energy and time invested in preparing the TWS. "(The TWS) took away time and energy from other responsibilities. It seems like busy work". . . ."The idea is great, but creating graphs can be time consuming." Forty-three of the 63 candidate responses to the second open-ended item asking for specific feedback on the seven TWS processes fell in the positive range. Candidates shifted their focus in this item to the benefits of the various TWS phases, but the underlying tone of time invested continued to surface. "Pieces were helpful (learning goals, assessment plan, analysis of student learning and reflection)". . . ."The phases were okay but (I) needed more time to implement (the TWS) in the classroom." The third opportunity for additional feedback demonstrated that candidates were clearly split in their overall perceptions of the TWS; 36 of the candidate comments fell into the positive range while 27 were scored as negative.

A second review of candidate comments on the open-ended items revealed that 12 of the respondents scored at the lowest or superficial level of Rodgers' reflective cycle. These candidates shared a negative view of this particular performance-based assessment. Eleven candidate respondents scored at the first level; these candidates primarily described their presence in the experiences. The second level, differentiating and looking at nuances, was reached by 22 of the respondents. Their comments touched on the benefits of the TWS assessment and on those pieces they felt were personally beneficial.

Fifteen of the participating candidates scored at the third level of the reflective cycle; their comments exemplified the ability to analyze the experience from multiple perspectives. Only three candidates responding to the open-ended items reached the highest level of Rodgers' reflective cycle. These candidates were able to share examples of learning to take intelligent action, testing ideas, and suggesting alternative ideas for future teaching and learning experiences.

Semi-Structured Interviews

This study was undertaken as a blended research design allowing the researchers to use both numbers and words to understand the perceptions of candidates and their mentors. This linking of quantitative and qualitative data facilitated a richness that expanded both the scope and breadth of the study. Semi-structured interviews were conducted with volunteers during a 30-90 day period after the end of the student teaching experience. Individual interviews were conducted either at the student teaching site, at the university, or by telephone in accordance with the preference of each participant. An analysis of the semi-structured interview transcripts revealed that the qualitative data supported the quantitative findings. In this limited sample, both candidates and their mentors expressed positive perceptions regarding the overall impact of the TWS on candidate performance. The few negative comments that were offered centered on work load issues and time constraints.

The semi-structured interview questions probed participants' perceptions regarding each of the seven TWS processes; in each case the responses were overwhelmingly positive and candidates scored consistently at level 2 or higher on the reflective cycle rubric. Mrs. A[1], a mentor, clearly expressed her positive perception of the contextual factors process. "She was so aware of their needs . . . and looked at them from a multicultural perspective . . . she did understand where people (parents) were coming from . . . adjusted expectations." Data from candidate interviews emphasized the benefits of using the contextual study to guide the development of learning goals. Mary Ellen, a candidate, commented, "It (the learning goals process) forced me to be organized and on task . . . and to really articulate the goals I wanted my students to do . . . (it) helped me to do a better job with the kids."

The comments of one candidate, Suzanne, exemplified the positive perceptions expressed by all of the interviewed candidates regarding the assessment planning process. "I would like to say, keep this part (assessment planning) . . . to me this was the most important part." and "Not only could I measure what my students were learning and how far they had come, but I could find out how far I had come!" In addition, the comments of candidate Katherine were representative of the overall positive view candidates expressed about the benefits of the instructional planning component of the Teacher Work Sample. "This design for instruction piece forces you to sit down and think about planning in a different way . . . as opposed to planning what kids might enjoy." and "This piece forces you to think about every single angle."

Mrs. A, a mentor, pointed out the benefits of asking candidates to reflect upon instructional decisions made in-action, "She (Mrs. A's candidate) reflected a lot, we would have big discussions and she would say . . . 'I know I should not have done that' and she would reflect with me frequently about her lessons." The researchers found that the impact of the analysis of the student learning process surfaced in the majority of semi-structured interview conversations. Jill, another candidate, explained her thoughts, "(Analysis is) probably

1 Pseudonyms have been assigned to all participants to guarantee privacy and confidentiality. Identifying characteristics have also been changed or omitted.

the most important part . . . because I had to know, did my students learn anything? . . . did I learn anything?"

Both candidates and mentors spoke about the reciprocal benefits of reflecting on teaching as well as reflecting on student learning. Carrie, a candidate, offered comments that summarize the overall perceptions expressed by participants, "Becoming a reflective practitioner . . . this piece truly helped me to do that. Now I am going to be able to go into this classroom in August and reflect on what I am doing every day!" All participants were in agreement that the TWS was helpful in developing the skills of a reflective practitioner. As Amy explained, "I think it (TWS) . . . made you sit down and actually think about what you do and need to do to be an effective teacher."

Discussion and Conclusions

Faculty members at Western Oregon University have been studying the perceptions of teacher candidates and in-service teachers regarding the effectiveness of the TWS for several decades (Girod & Shalock, 2002). The two most common views expressed by Western Oregon candidates were that the TWS was a key factor in helping them to become very focused as teachers and that the process of developing a TWS deepened their understanding of complex instructional units (Girod & Shalock, 2002). California teacher educators involved with the PACT performance-based assessment have experienced similar findings. While their candidates agree that the process is time consuming, many state that they have learned a great deal from the experience (Rothman, 2008). Data gathered from candidate participants of this study seem to concur with the findings at Western Oregon and the California consortium of universities. It appears that the concepts and skills nurtured by the TWS performance-based assessment are not all that different from the concepts and skills traditionally taught in teacher preparation programs. The benefit appears to be that the design and implementation of a comprehensive TWS unit allows candidates to take the theory emphasized in university classrooms and effectively apply it in the authentic classroom setting. Another significant feature of the TWS is that the processes facilitate candidate analysis and reflection regarding group and individual student progress (Delvin-Scherer, Daly, Burroughs, & McCartan, 2007).

Mentors' perceptions regarding the impact of the TWS on raising the quality of candidates' instructional effectiveness were generally more positive than those of candidates. Furthermore, mentor responses to the open-ended items and semi-structured interview probes supported the positive responses gathered via the survey instruments. Subsequent data analyses revealed patterns consistent with candidate concerns regarding time and workload issues caused by the TWS during the student teaching experience. Many expressed the difficulty of supporting candidates as they juggled this added task to an already burdensome workload. In this study, elementary education candidates often struggled to make developmentally appropriate decisions about teaching and learning during the student teaching experience; at the same time candidates were trying to complete the numerous TWS requirements.

Although both candidates and mentors demonstrated positive perceptions regarding the impact of the TWS on raising the quality of candidates' instructional effectiveness, there were significant differences in their responses to 11 out of 17 items on the survey instrument. For these 11 items, mentors perceived the impact of the TWS more positively than did candidates. This difference may be attributed to the fact that mentors were allowed the luxury of observing candidates' as they grew in competence; candidates themselves may have developed more positive perceptions given time and distance from the experience for thought and reflection.

A second theme emerged from the mentor responses in both open-ended items and semi-structured interviews: mentors expressed a need for clarity regarding their own roles as mentors for candidates throughout this performance-based assessment. Similar findings were reported by TWS researchers at Seton Hall University. Mentors participating in that study also requested clear guidelines and suggestions for better support of their candidates in the design and implementation of a TWS (Devlin-Scherer, Daly, Burroughs, & McCartan, 2007).

In terms of candidate qualitative responses, the majority of these were at the lower levels (i. e., Level 0 & Level 1) on the reflective cycle rubric. A potential contributing factor to candidate responses here may have been one of timing. The survey instruments were administered during the final class meeting of the student teaching semester. Typically, any candidate at this point in the academic term would be struggling to balance numerous academic and personal responsibilities. Candidates at this university may have responded differently if more time between the experience and the administration of the survey instrument had allowed for depth in reflection.

The majority of semi-structured interview participants met with the researcher approximately 30 days after the completion of a 16-week student teaching experience. Interestingly, semi-structured interviews with candidates revealed overwhelmingly positive perceptions regarding the TWS. Candidate responses in these dialogues consistently fell into the upper levels (i. e., levels 2 – 4) of the reflective cycle rubric. Time and distance from the experience may have been a contributing factor in the positive responses offered during semi-structured interviews. In fact, candidates who were interviewed 30-90 days after the completion of the student teaching experience expressed more positive overall views and reached the highest levels on the reflective cycle rubric.

The face-to-face conversations may have been another factor that encouraged candidates to move further into the reflective cycle. During the semi-structured interviews, the researcher had the opportunity to observe the level of reflective thinking change as candidates made connections,

reorganized, and reconstructed their experiences in community (Rodgers, 2002b). As the interview conversations progressed the candidates described and differentiated the experiences of designing and implementing a TWS. These reflective conversations allowed the candidates to slow down and examine the various nuances of their own experiences (Spelman & Allman, 2007).

Interestingly, similar findings were discovered by a team of researchers piloting the TWS as part of the student teaching experience at the University of Northern Iowa. Mentors in this pilot program agreed that the TWS was effective and served to better structure the student teaching experience. However, they agreed with the mentors participating in this study as they cautioned that careful management of the overall workload was critical to the success of the TWS experience (Henning, DeBruin-Parecki, Hawbecker, Nielsen, Joram, & Gabriele, 2005).

Researchers at Western Oregon University found that implementing TWS Methodology is very contextual and that each program faces unique challenges as they implement TWS assessments (Wright, 2002). Several other studies echo the findings of this study leading to the conclusion that reflective practice is an important outcome for those programs hoping to prepare highly-qualified teachers. Researchers at Seton Hall University agreed that the TWS experience resulted in far more compelling reflection pieces than previous practices (Devlin-Scherer, Daly, Burroughs, & McCartan, 2007). Similar reports were cited in the Northern Iowa University pilot program. Researchers there noted that the TWS proved to be a beneficial tool for promoting reflection on student learning (Henning, DeBruin-Parecki, Hawbaker, Nielsen, Joram, & Gabriele, 2005). The Interstate New Teacher Assessment and Support Consortium's (INTASC, 1991) ninth core standard for teachers states: "The teacher is a reflective practitioner who continually evaluates the effect of his/her choices and actions on others (students, parents, and other professionals in the learning community) and who actively seeks out opportunities to grow professionally." Effective teachers, then, engage in reflective practice.

Lessons Learned

Mewborn (1999) argued that pre-service teachers need time to learn and practice reflective skills in a non-evaluative environment. It is essential then that teacher educators support the growth of reflective skills throughout the pre-service and in-service development of professional educators. Thus, it will be important that the elementary education faculty at this university find ways to scaffold candidates' reflective habits. Ensuring time and distance from the experiences as well as providing opportunities for reflection in community will need to be added to early professional education coursework as well as to the student teaching experience. In fact, Bullough & Baughman (1997) asserted that the first five to seven years of teaching careers constitute the novice period; these years should be marked by ongoing reflection. These arguments provide food for thought as the researchers in this study review, reflect, and revise TWS performance-based assessment for future elementary education program candidates.

The TWS at this university should be further contextualized to provide a means of addressing not only the need for measurement and accountability, but to serve as an opportunity to strengthen the reflective skills of candidates. Candidates, mentors, and all program faculty should be introduced to a common set of reflective stages and embed one vehicle for assessing candidate reflective growth throughout professional education coursework.

Overall the TWS has been a positive addition to this particular elementary education teacher preparation program. However, based on the pilot study, the researchers learned a few lessons that may inform other institutions. First, embedding the seven processes of the TWS backwards into professional coursework may increase the comfort level of candidates implementing a TWS in the final student teaching experience. Early survey courses may include contextual studies of field experience sites; learning theory courses could require candidates to develop learning goals and align those goals with state standards. Various methods courses can introduce candidates to the assessment planning and instructional design processes. Later coursework may require candidates to analyze assessment data and plan a mini-TWS. The TWS would then become the common thread that weaves teacher preparation coursework into a comprehensive tapestry of preparation. In addition, these steps may help to raise candidates' perceptions regarding the impact of the TWS on their own teaching effectiveness.

Next, a TWS handbook to guide and inform candidates could be developed. Such a handbook could provide timelines, worksheets, graphic organizers, and scoring rubrics for each of the seven distinct processes. Candidates could use a contextualized handbook to guide the development and implementation of the TWS during the student teaching experience. In addition, breaking the assignment into seven separate pieces would allow for formative feedback, collegial conversations, peer evaluations, and opportunities for revision and resubmission. Another support for future candidates might be the sharing of exemplars and non-exemplars to strengthen their vision of effective TWS design.

Last, orientation and support for mentors should be a major part of TWS implementation. Orientation sessions and workshops for mentors could help to build a common understanding and language that can be useful in supporting candidates in the TWS design and implementation. Such meetings might also open the lines of communication between universities and P-12 field experience schools. A guidebook for mentors should also be developed. A streamlined guidebook for mentors might address common issues, define terms, offer examples, and even provide answers for frequently asked questions.

Implications for Future Research

Subsequent studies should investigate opportunities for purposeful pairings of candidates and mentors to facilitate successful student teaching experiences (Spelman & Allman, 2007). In addition, it is important to explore the advantages and disadvantages of placing candidates with program alumni or current graduate students more familiar with the TWS Methodology.

Finally, the nature of reflective growth in both pre-service and in-service teachers demonstrates the need for ongoing research regarding the effectiveness of the TWS Methodologies as they relate to reflective practitioners. There is a need for longitudinal studies that follow teacher candidates' progress as they move into the role of professional educators. In addition, the researchers see a need for continued gathering of stakeholder perceptions as well as ongoing program evaluation designed to inform future program improvement and related modifications of the TWS.

References

Bullough, R., & Baughman, K. (1997). *First year teacher eight years later: An inquiry into teacher development*. New York: Teachers College Press.

Cochran-Smith, M. (2005). No child left behind: 3 years and counting. *Journal of Teacher Education*, *56*, 99-103.

Cochran-Smith, M., & Zeichner, K., (Eds.). (2005). *Studying teacher education: The report of the AERA panel on research and teacher education.* Washington, DC: American Educational Research Association.

Cuban, L. (1990). Reforming again, again, and again. *Educational Researcher*, *19*, 3-13.

Darling-Hammond, L., & Ball, D. (2002). *Teaching for high standards: What policymakers need to know and be able to do*. Philadelphia: Consortium for Policy Research in Education.

Delvin-Scherer, R., Daly, J., Burroughs, G., & McCartan, W. (2007). The value of the Teacher Work Sample for improving instruction and program. *Action in Teacher Education*, *29*(1), 51-60.

Denner, P., Norman, A., Salzman, S., & Pankratz, R. (2003, February). *Connecting teaching performance to student achievement" A generalizability and validity study of the Renaissance Teacher Work Sample Assessment.* Paper presented at the annual meeting of the Association of Teacher Educators. Jacksonville, FL.

Dewey, J. (1933). *How we think.* Buffalo, NY: Prometheus Books. (Original work published in 1910).

Girod, G. (Ed.). (2002). *Connecting teaching and learning: A handbook for teacher educators on teacher work sample methodology.* New York: AACTE Publications.

Girod, M., & Girod, G. (2006). Exploring the efficacy of Cook School district simulation. *Journal of Teacher Education*, 57, 481-497.

Girod, G., & Shalock, M. (2002). Does TWSM work? In G. Girod (Ed.), *Connecting teaching and learning: A handbook for teacher educators on teacher work sample methodology (*pp. 347-358*)*. New York: AACTE Publications.

Greenwald, R., Hedges, L., & Laine, R. (1996). The effect of school resources on student achievement. *Review of Educational Research*, 66, 361-396.

Hamel, F., & Merz, C. (2005). Reframing accountability: A preservice program wrestles with mandated reform. *Journal of Teacher Education*, 56, 157-167.

Henning, J., DeBruin-Parecki, A., Hawbaker, B., Nielsen, C., Joram, E., & Gabriele, A. (2005). The teacher work sample: A tool for scaffolding and assessing preservice teachers' early field experiences. *The Teacher Educator*, *40*(3), 188-207.

INTASC (1991). *Model standards for beginning teacher licensing and development, Interstate New Teacher Assessment and Support Consortium*. Washington, DC: Council of Chief State School Officers.

Ingersoll, R. M. (1996). Out-of-field teaching and educational equality. Washington, DC: National Center for Educational Statistics.

Mewborn, D. (1999). Reflective thinking among preservice elementary mathematics teachers. *Journal for Research in Mathematics Education*, *30*, 316-341.

National Board for Professional Teaching Standards (NBPTS). *Toward high and rigorous standards for the teaching profession.* Detroit: NBPTS, 1989.

National Commission on Teaching & America's Future. (2003). *No dream denied: A pledge to America's children.* New York: National Commission on Teaching & America's Future.

National Commission on Teaching & America's Future. (1996). *What matters most: Teachers for America's future.* New York: National Commission on Teaching & America's Future.

Patton, M. (2002). *Qualitative research and evaluation methods.* Thousand Oaks, CA: Sage Publications.

Pankratz, R. (1999). *Improving teacher quality through partnerships that connect teacher performance to student learning.* Unpublished manuscript, Western Kentucky University.

Rodgers, C. (2002a). Defining reflection: Another look at John Dewey and reflective thinking. *Teachers College Record*, *104*, 842-865.

Rodgers, C. (2002b). Seeing student learning: Teacher change and the role of reflection. *Harvard Educational Review*, *72*, 230-253.

Rossman, G., & Rallis, S. (2003). *Learning in the field: An introduction to qualitative research.* Thousand Oaks, CA: Sage Publications.

Rothman, R. (2008). Taking the measure of new teachers. *Harvard Education Letter*, *24*(4), 1-3.

Schalock, H., & Myton, D. (2002). Connecting teaching and learning: An introduction to teacher work sample methodology. In G. Girod (Ed.), *Connecting teaching and learning: A handbook for teacher educators on teacher work sample methodology* (pp. 5-31*)*. New York: AACTE Publications.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*, 1-22.

Spelman, M. V. (2006). From policy to practice: A case study of initial certification reforms in schools of education. *Dissertation Abstracts International*, 67(02), (UMI NO. 3202747).

Spelman, M. & Allman, T. (2007). Coaching & mentoring in literacy field experiences: A comparison between traditional and purposeful mentoring pairings. *The Reading Professor 29*(1), 22-29.

Wise, A. E., & Leibbrand, J. (1996, November). Profession-based accreditation: A foundation for high-quality teaching. *Phi Delta Kappan*, *78*, 202-207.

Wright, D. (2002). Teacher work sample methodology in early childhood and elementary preparation: A case study. In G. Girod (Ed.), *Connecting teaching and learning: A handbook for teacher educators on teacher work sample methodology* (pp. 361-374*)*. New York: AACTE Publications.

Zumwalt, K. & Craig, E. (2005). Teacher characteristics: Research on the indicators of quality. In M. Cochran-Smith, & K. Zeichner, (Eds.), *Studying teacher education: The report of the AERA panel on research and teacher education* (pp. 111-156*)*. Washington, DC: American Educational Research Association.

David Bell specializes in curriculum and instruction, school-wide diversity transformation, and closing the achievement gap for K-8 students in high needs schools. E-mail: bell@sxu.edu.

Maureen Spelman research interests include teacher preparation, teacher work sample methodologies, closing achievement gap for K-8 students in high-needs schools, and the impact of educational reforms on schools of education.

Holly Mackley specializes in literacy and literature for early childhood and elementary education.

Liang Zhao research interests are in the history and philosophy of education, technology in education, and research methodology.

Assessment Systems and Data Management in Colleges of Education: An Examination of Systems and Infrastructure

Noela A. Haughton
Virginia L. Keil
The University of Toledo

Abstract

The College of Education Assessment Infrastructure Survey was developed and administered to 1011 institutions over a twelve-month period ending April 2007. The survey examined the capacity of university-based teacher preparation programs to respond to the growing and increasingly complex data management requirements that accompanies assessment and accountability edicts. The summarized responses of 266 institutions (26%) confirmed that education units have responded to accreditation and accountability mandates though none have relied solely on the institutional-level infrastructure. This has resulted in the implementation of a variety of assessment systems that vary widely in sophistication and cost. Also, the management of these systems is largely a part-time endeavor. Additional and more contextually based studies must be conducted to determine the actual level and type of institutional support, other hidden costs, and the level of return on investment in terms of the effectiveness of these systems and their impact on unit quality outcomes.

Introduction

Higher education in the United States has experienced unprecedented growth in demand in recent decades. According to the National Center for Education Statistics (n.d.), there was a 16% increase in enrollment at degree-granting institutions between 1985 and 1995 and a 23% increase between 1995 and 2005. This expansion has been fueled by a number of factors among which is online and distance education which has lead to an increasingly competitive market. This in turn has led to significant investments in information and educational technology infrastructure, including hardware and software, staffing, and funding levels.

In 2000-01, 56% of all degree-granting institutions offered distance courses, offering 127, 400 online courses to 3,077,000 students (Bradburn & Zimbler, 2002). Allen and Seaman (2007) reported that higher education institutions taught nearly 3.2 million students online in the fall 2005 semester. This represented an approximately 35% growth over the previous year. It is reasonable to expect that the increasing enrollment trend will continue along with infrastructure expenditure. Hawkins and Rudy (2007) reported that for the 2005–2006 fiscal year, the median operating appropriation to centralized information technology (IT) functions for doctoral, master, and bachelor's level institutions were $10.896 million, $2.450 million, and 1.488 million respectively. The top 16 areas of expenditures include network infrastructure and services, web support, research computing, distance education, and instructional technology. In terms of staff, this function was mainly supported by fulltime equivalent (FTE) regular employees though there was some reliance on student FTE employees. The highest percent of staff, across all Carnegie levels, were assigned to administrative/enterprise information systems, and the functional area that includes desktop computing support, user support services, and training. Many institutions realize that this level of investment in technology and support infrastructure is necessary to support strategic endeavors and enterprise-wide data management and thus their ability to compete in an increasingly competitive and global market. Trends indicate that higher education will continue to grow and diversify, which in turn will continue to drive investment in technology and support infrastructure at the institutional level.

This growth in higher education demand has also been accompanied by calls for increased accountability especially related to learning outcomes. Since the Nation at Risk report in 1982, eyes have been on the effectiveness of higher education. It is no longer enough to measure effectiveness in terms of inputs such as enrollment, degrees granted, size of library holdings, etc. Rather, institutions are being required by accreditors and other entities to use outcomes-based evidence to demonstrate knowledge, skills, and workforce readiness. However, it is unclear if existing data management and reporting infrastructures within higher education are appropriate for the new type of outcomes assessments. Along with these assessments come the evidence gathering, data management, analysis, and reporting required to demonstrate successful outcomes. Many would agree that a significant, and some would argue disproportionate, amount of the accountability pressure in higher education is focused upon schools, colleges, and departments of education (SCDE), who prepare the majority of new teachers for the nation's elementary and secondary schools.

Assessment, Accreditation, and Accountability

Palomba and Banta (1999) define assessment as "the systematic collection, review, and use of information about educational programs undertaken for the purpose of improving student learning and development" (p. 4). For SCDE, this system of collection, review, and use has become increasingly prescriptive and now includes documenting the impact of teacher candidates on the learning of elementary and secondary school students (Diez, 2001). While the Teacher Education Accreditation Council (TEAC) quality principles do not make specific reference to a technology-based assessment system, programs are required to use multiple measures and assessment methods to demonstrate pre-service teacher learning. The National Council for Accreditation of Teacher Education (NCATE) requires an assessment system "that collects and analyzes data on applicant qualifications, candidate and graduate performance, and unit operations to evaluate and improve the performance of candidates, the unit, and its programs" (NCATE, 2008, p. 25). It is fair to say that the assessment requirements of both accreditors focus on the use of multiple methods and ongoing feedback for the candidates and ongoing diagnosis of program effectiveness for the education unit. Such a system will necessarily rely on the support of an infrastructure of technology, staffing, and most importantly, initial and ongoing funding to effectively support both accountability strands.

While there is some state and federal support for some assessment systems, funding remains largely the responsibility of the institution (Carey, 2007). In the age of shrinking budgets, the appropriate funding level of an assessment system may be impacted by the reality of the needs of other critical college and institutional demands. Regardless, the use of technology-supported assessment systems to document learning outcomes and support continuous improvement is mandated by unit accreditors (NCATE, 2008; TEAC, 2006). The degree to which SCDE have responded to the infrastructural demands is unclear. What is increasingly clear is, given the current climate in which SCDE must compete with other strategic initiatives such as distance education, many may be struggling to respond to this high stakes mandate.

The vast diversity of SCDE raises another issue—consensus regarding minimum expectations of the infrastructure in terms of personnel, technology, etc., of these assessment systems. It is reasonable to expect that the diversity of SCDE in terms of unit size, Carnegie classification, number of graduates, and size of faculty would influence the assessment, data management, and reporting needs. Also, it is very likely that this diversity has contributed to the lack of consensus regarding the minimum configuration of these systems. This, along with some of the previously mentioned constraints, has resulted in SCDE's relative autonomy regarding implementing contextually appropriate assessment systems. These range in sophistication from completely paper-based systems to proprietary in-house-developed solutions (i.e., a complete assessment system designed and developed in-house using a programming language such as ASP, Visual Basic, C++, etc.). Lastly, this diversity in SCDE may influence the personnel support provided to the assessment function. This support includes the coordinator's role and responsibilities, the degree to which this role is designated full-time or part-time, and the number of personnel assigned to the assessment function on a continuous basis.

Irrespective of assessment system choice and operating context, a well-designed system requires significant investment infrastructure, including human, technology, and other resources. Also, such a solution, regardless of unit characteristics, should effectively support the assessment and reporting function that is critical to meeting accreditation and accountability mandates.

This is the first in a series of articles that will focus on how SDCE are meeting accountability mandates. Specific questions addressed are:

1. What type of assessment systems are SCDE implementing?
2. Who coordinates the unit's assessment function?
3. What workload arrangements are in place for the unit's assessment coordinator?
4. How many people are assigned to support the assessment function?
5. What is the estimated annual cost of the assessment function?

Method

Participants

The College of Education Assessment Infrastructure Survey (CEAIS) was administered to 1,011 schools, colleges, and departments of education unit heads over a 12-month period ending April 2007. The lack of ready access to a comprehensive and accurate national list of university-based education programs, which included unit heads along with their email addresses, led the researchers to compile a list of colleges of education from the US News and World Report (search.usnews.com) college finder. In February 2006, a list of 1,152 institutions that offered teaching preparation in any field was created.

From this list, the names of unit heads—including deans, department chairs, and program directors—were compiled from the school web sites, which were hyperlinked to the institution names captured from the US News web site. Locating the contact information of these institutions was a time-consuming process. Therefore, to facilitate a fairly quick pilot study, a sub-list of 230 institutions from the Council of Academic Deans from Research Education Institutions (CADREI), The Renaissance Group, and the Texas Association of Colleges for Teacher Education (TACTE) was created. The first electronic administration to this sub-list took place over a five-month period from April 2006 to September 2006. Valid contact information was located for an additional 781 institutions to which the second administration took place

from November 2006 through April 2007. The targeted total response rate was 279 institutions, which is required for a 95% confidence level.

Instrumentation and Procedures

The CEAIS was developed by the authors to determine assessment infrastructure and practices of units of education. Items for this questionnaire came from multiple sources. As previously mentioned, units of education are diverse in terms of a number of characteristics including accreditation status, Carnegie classification, number of graduates, whether they are public or private, and the availability of resources. Despite this diversity, all assessment systems are required to accomplish specific goals. These are: 1) collect performance evidence from a minimum number of sources; 2) aggregate and disaggregate data on key characteristics such as licensure area and key student characteristics including gender and licensure route; 3) facilitate the sharing of data with and reporting performance information to critical stakeholders including teacher education faculty, candidates, and community partners; and 4) support data-driven changes and decisions to continuously improve programs and teacher quality. Accordingly, the CEAIS was designed to capture data surrounding these issues and requirements.

Content validation of the questionnaire was undertaken throughout its development. The primary methods were alignment with accreditation requirements and expert-paneling by faculty from different program areas including some with accreditation-related responsibilities. Prior to each electronic administration, the questionnaire was tested for flow, readability, appearance, and technical bugs by the researchers, and other faculty and staff colleagues.

At first administration, this instrument was a 32-question, seven-page, web-based survey designed to solicit responses regarding the assessment practices in university-based teacher education programs.

- About the Institution (Page 1) had five questions that solicited demographics data including: respondent position, accreditation status, Carnegie classification, and institution affiliation (CADREI, Renaissance Group, or TACTE).
- About the Unit of Education (Page 2) had six questions that solicited unit characteristics in terms of: accreditation status; accrediting body; program size in terms of graduates; and program size in terms of faculty—fulltime and part-time.
- About the Assessment System (Page 3) had four questions that solicited unit assessment system data, including: status of system (e.g., conversion vs. full implementation), type of electronic system (e.g., paper-based, proprietary, supported by productivity software, and supported by third party software), and reasons for assessment system choice (cost, ease of implementation, product support, popularity with other units of education, etc.).
- About the Assessment Support Structure (Page 4) had five questions related to personnel support structure, including: who—faculty or staff –coordinates assessment function, workload arrangements for the coordinator of the assessment function, number of staff assigned to the assessment function on a fulltime basis, number of staff assigned to the assessment function on a part-time basis, and estimated total annual expenditure on assessments.
- About Experience with Sharing Candidate Data (Page 5) had two questions related to experience with sharing candidate data with faculty including: if data was shared, and the type of data (performance on licensure tests, candidate ability to plan instruction, etc.).
- About Faculty Receptivity towards Assessment Data (Page 6) had seven questions solicited information regarding faculty receptivity towards each of the different type of assessment data, e.g., performance on licensure tests, ability to plan instruction, etc.
- Questions About Experience with Sharing Candidate Data (Page 7) had three additional questions related to sharing data including: what accounts for the type of receptivity, whether changes were made as a result of what was learned from the data, and the details regarding the type of changes made.

It was initially anticipated that the first administration period would be a month. However, the low return—40 responses—resulted in a change of strategy to extend the collection period until September when most schools were back in session. Re-dropping (Alreck & Settle, 2004) was done via electronic mail to non-responding institutions. This excluded those who specifically declined to participate in the study. Improvements to the CEAIS were made based on feedback from the initial administration. Specific changes were:

- About the Institution (Page 1): no changes. Data included in this analysis.
- About the Unit of Education (Page 2): no changes. Data included in this analysis.
- About the Assessment System (Page 3): no changes. Data included in this analysis.
- Assessment Support Structure (Page 4): a new choice—"ability to aggregate and disaggregate data"—was added as a response choice to the "reason for assessment system choice." This function is critical to the analysis needed to demonstrate accountability. Therefore, despite potential internal validity concerns, the researchers decided to include this option because of its potential to improve the survey. Also, three questions—assessment coordinator gender, highest degree earned, and prior assessment experience—were added. However, these additional questions were not included in this analysis.
- Experience with Sharing Candidate Data (Page 5): two questions related to the sharing of data with Arts and Science faculty and stakeholders impacted by data-driven

changes were added. However, questions on this page of the survey were not included in the analysis.

- Faculty Receptivity towards Assessment Data (Page 6): two questions related to the receptivity of Arts and Science faculty and other stakeholders were added. However, questions on this page of the survey were not included in this analysis.
- Experience with Sharing Candidate Data (Page 7): no changes.

The second administration took place from November 2006 through April 2007 and included the remaining 781 institutions. Once again, a low response rate—98 responses—prompted an electronic reminder, which, once again, excluded respondents who declined to participate, of which there was a total 86 institutions across both electronic administrations. At the end of the electronic data collection period, another sub-set of 63 large, mostly public, non-responding intuitions were identified for a third paper and pencil administration based on the updated form. No follow-up was done for this final administration.

Data Analysis

Web-based results from both electronic administrations were exported from Survey Monkey into a text file that was subsequently imported to SPSS (13) for Windows. The nine paper-based survey responses were added manually and coded separately. Responses were coded and a series of summaries and cross-tabulations based on key institutional characteristics (e.g., Carnegie classification and institution size) were produced for each of the five questions being addressed in this report. Percentages of respondents (column totals) based on these characteristics were also reported in parentheses.

Results

Description of Responding Units

Seventy-one responses (approximately 31% of the 230 surveyed institutions) from administration one, 186 responses (approximately 24%) from administration two, and nine responses (approximately 14%) from administration three were combined for a total of 266 responses—a combined 26% response rate—were analyzed. While this sample size is well above the 214 required for a 90% confidence level for a target population of 1,011, it is 13 short of the 279 required for a 95% confidence level. The respondents included: 177 (66.5%) deans or unit heads; 27 (10.2%) associate deans; 29 (10.9%) faculty members, one of whom had assessment coordinator duties; 23 (8.7%) administrators who are assessment coordinators or assessment directors; and 10 (3.8%) staff members.

Table 1 describes the diverse sample of responding institutions in terms of institutional and unit characteristics. Fifty percent (133) of the reporting institutions are public with the remaining half private. In terms of regional accreditation: 42 (15.8%) are from the Middle States Association of Colleges and Schools, 18 (6.8%) from the New England Association of Schools and Colleges; 91 (34.2%) from North Central Association of Colleges and Schools, 13 (4.9%) from the Northwest Commission on Colleges and Universities, 10 (3.8%) the Western Association of Schools and Colleges, and 89 (33.5%) from the Southern Association of Colleges and Schools. In terms of professional (unit) accreditation, 170 are accredited while 65 are not. Of the professionally accredited units: 152 (89.4%) are NCATE accredited, 17 (10%) are TEAC accredited, and one (0.6%) is accredited by NCATE and TEAC.

Unit size is categorized by four indicators: Carnegie classification, average number of graduates (three years), number of full-time faculty, and number of part-time faculty. The majority of the reporting institutions—120 (45.1%)—are master's level, 75 (28.2%) are doctoral granting, and 70 (26.3%) are bachelor's level. One hundred and five (39.5%) graduate less than 100 new teachers annually, 63 (23.7%) graduate between 100 and 249, and the remaining 81 (30.5%) programs graduate at least 250 new teachers. In terms of number of full-time faculty: the majority of the reporting institutions—143 (53.8%)—employ fewer than 25 faculty members, 68 (25.6%) between 25 and 74, and 39 (14.7%) employ 75 or more. The employment figures for part-time faculty generally reflect those of full-time faculty. One hundred and fifty-one (56.8%) respondents report fewer than 25, 61 (22.9%) employ between 25 and 74, and 26 (9.8%) employ 75 or more. These characteristics indicate that the responding units represent a cross-section of preparation programs, adding credibility to the results.

What type of assessment systems are SCDE implementing?

Table 2 summarizes the responses to this question and supports the notion that SCDE have not been able to rely on the larger institutions' systems to respond to assessment and accountability requirements. This in turn has led to a very vast diversity of assessment systems. Thirty (12.6%) units designed and developed proprietary assessment systems, 93 (38.9%) implemented assessment systems supported with productivity software, 91 (38.1%) implemented assessment systems supported with third-party assessment software, and 22 (9.2%) used assessment systems that were primarily paper-based/manual. In this sample, units are most likely to use assessment systems either supported with productivity software (e.g., Microsoft Office, FileMaker Pro, etc.) or third-party assessment software (e.g., Livetext, Task Stream, etc.).

The strong presence of a third party assessment system software market is further evidence that existing reporting systems in higher education, in and of themselves, are not sufficient to support SCDE's reporting requirements. More than one third of the responding units use these systems. The three most commonly used third-party assessment system packages are Livetext (39 users), Task Stream (14 users), and

Table 1
Responding Institutions' Descriptive Statistics

Public / Private (266 responses)		Unit Accreditation (235 responses)			
Public: 133 (50%)	Private: 133 (50%)	NCATE: 152 (57%)	TEAC: 17 (6.4%)	NCATE & TEAC: 1(.4%)	Not Accredited: 65 (24.4%)

Regional Accreditors (263 responses)					
Middle States: 42 (15.8%)	New England Assoc: 18 (6.8)	North Central: 91 (34.2%)	Northwest Commission: 13 (4.9%)	Western Association: 10 (3.8%)	Southern Association: 89 (33.5%)

Carnegie Classification (265 responses)		
Bachelors: 70 (26.3%)	Masters 120 (45.1%)	Doctoral: 75 (28.2%)

Number of Graduates (249 responses)							
< 50: 57 (21.4%)	50 – 99: 48 (18%)	100 – 149: 30 (11.3%)	150 – 199: 19 (7.1%)	200 – 249: 14 (5.3%)	250 – 299: 11 (4.1%)	300 – 349: 14 (5.3%)	> 350 or more: 56 (21.1%)
< 100: 105 (39.5%)		100 – 249: 63 (23.7%)			250 or more: 81 (30.5%)		

Unit Size: Fulltime Faculty (250 responses)									
< 10: 76 (28.6%)	10 – 24: 67 (25.2%)	25 – 49: 41 (15.4%)	50 – 74: 27 (10.2%)	75 – 99: 22 (8.3%)	100 – 124: 11 (4.1%)	125 – 149: 0 (0%)	150 – 175: 4 (1.5%)	175 – 199: 0 (0%)	200 or more: 2 (.8%)
< 25: 143 (53.8%)		25 – 74: 68 (25.6%)		75 or more: 39 (14.7%)					

Unit Size: Part-Time Faculty (238 responses)									
< 10: 88 (33.1%)	10 – 24: 63 (23.7%)	25 – 49: 39 (14.7%)	50 – 74: 22 (8.3%)	75 – 99: 9 (3.4%)	100 – 124: 14 (5.3%)	125 – 149: 0 (0%)	150 – 175: 0 (0%)	175 – 199: 2 (.8%)	200 or more: 1 (.4%)
< 25: 151 (56.8%)		25 – 74: 61 (22.9%)		75 or more: 26 (9.8%)					

TK20 (11 users). Paper-based systems are more likely to be used by un-accredited units at private master's level institutions. The reasons for assessment system choice indicate that there is an effort and a need to balance accountability expectations with the reality of resource availability. The top three reasons for assessment system choice are ease of use (122 responses), cost (119 responses), and the ability to customize (102 responses). The top two reasons indicate that access to the resources needed to effectively implement and manage these systems is an issue for many SCDE. While "The ability to aggregate and disaggregate data" was not available as a choice in the first version of the survey primarily due to oversight, it is noteworthy that this reason is the second choice of the 195 respondents who had this option.

Who coordinates the unit's assessment function?

Table 3 summarizes assessment coordinator roles by institutional and unit characteristics. Results indicate that the assessment and data management function is managed by faculty in many units. A small majority—120 (51.18%) of the 235 responding institutions—report that their assessment coordinators are faculty members. Of these faculty members: 44 (16.5%) are untenured, 66 (24.8%) are tenured, and 10 (2.3%) are non tenure track. Other coordinators include non-faculty administrators or staff: 100 (37.6% of assessment coordinators), nine (3.4%) mixed teams of faculty and/or non-faculty, and six (2.3%) other coordinating arrangements (e.g., consultants).

In cases where faculty members coordinate the assessment function, untenured faculty are employed in this role in significant numbers: 22 (18.3%) in public institutions and 22 (19.1%) in private institutions; 14 (23%) in bachelor's level, 15 (14.2%) in masters granting institutions, and 15 (22.4%) doctorate granting institutions; 34 (21.3%) of professionally accredited units and 9 (15%) of un-accredited units; 19 (19.8%) smaller (number of graduate) programs, 11 (17.7%) medium sized programs, and 14 (18.7%) of larger programs. It is unclear whether a faculty vs. a non-faculty coordinator presents an advantage or a challenge. However, some may argue that untenured faculty in this role potentially face two competing, and arguably at research universities, un-related endeavors: the challenge of earning tenure and the challenge of managing the unit's assessment and accountability requirements.

What workload arrangements are in place for the unit's assessment coordinator?

Table 4 summarizes workload arrangements by institutional and unit characteristics, and coordinator role, respectively. Of the 223 responses, the overwhelming majority—181 (81.2%)—indicate that the assessment coordinator is mostly a part-time role. Part-time describes any non-full-time appointment, including extra compensation. Sixty-two (23.3%) report that assessment duties are part-time based on release time from other duties; 99 (37.2%) respondents report part-time duties of the position; six (2.3%) are based on extra compensation; and eight (3%) are based on overload or added job responsibilities. Twelve (18.5%) respondents report full-time duties and five (7.7%) indicate other arrangements. Also, in this sample, faculty coordinators are less likely to be full-time (13.6%) than non-faculty (23.9%). This reliance on part-time effort supports the notion that resource allocation to the assessment and accountability effort continues to be an issue for many SCDE.

How many people are assigned to support the assessment function?

Table 5 summarizes the number of fulltime and part-time personnel support for 233 and 230 responding institutions respectively. An overwhelming majority of units—215 (80.8%)—have fewer than three persons assigned to the assessment function on a fulltime basis. Fifteen (5.6%) and

Table 2
Type of Electronic Data Management Systems by Institution and Unit Characteristics

		Public/Private (239 responses)		Carnegie (238 responses)			Unit Accreditation (225 responses)		# Graduates (237 responses)		
Institution Characteristics	All	Public	Private	Bach	Mast	Doc	Yes	No	< < 100	100–249	250 or more
Proprietary	30 12.6%	23 18.9%	7 6%	6 9.5%	15 13.6%	8 12.3%	26 16%	3 4.8%	8 8.1%	6 9.7%	16 21.1%
In-house (MS Office, SPSS, etc.)	93 38.9%	39 32%	54 46.2%	32 50.8%	43 39.1%	18 27.7%	63 38.7%	24 38.7%	48 48.5%	25 40.3%	20 26.3%
Supported with third-party system	91 38.1%	55 45.1%	36 30.8%	18 28.6%	37 33.6%	36 55.4%	67 29.8%	19 30.6%	31 31.1%	24 38.7%	34 44.7%
Primarily paper-based	22 9.2%	3 2.5%	19 16.2%	7 11.1%	13 11.8%	2 3.1%	6 3.7%	16 25.8%	11 11.1%	5 8.1%	6 7.9%
Other	3 1.3%	2 1.6%	1 .9%		2 1.8%	1 1.5%	1 .6%		1 1.0%	2 3.2%	
# Responses	239 100%	122 100%	117 100%	63 100%	110 100%	65 100%	163 100%	62 100%	99 100%	62 100%	76 100%

Table 3
Assessment Coordinators by Institution and Unit Characteristics

Position description	All	Public/Private (235 responses)		Carnegie (234 responses)			Unit Accreditation (220 responses)		# Graduates (233 responses)		
		Public	Private	Bach	Mast	Doc	Yes	No	< < 100	100–249	250 or more
Faculty—untenured	44 16.5%	22 18.3%	22 19.1%	14 23%	15 14.2%	15 22.4%	34 21.3%	9 15%	19 19.8%	11 17.7%	14 18.7%
Faculty—tenured	66 24.8%	35 29.2%	31 27%	19 31.1%	28 26.4%	18 26.9%	49 30.6%	12 20%	26 27.1%	12 19.4%	28 37.3%
Faculty—non tenure track	10 2.3%	5 4.2%	5 4.3%	3 4.9%	3 2.8%	4 6%	7 4.4%	2 3.3%	5 5.2%	2 3.2%	2 2.7%
				Faculty Only Coordinators							
	120 51.1%	62 51.7%	58 50.4%	36 59%	46 43.4%	37 55.2%	90 56.2%	23 38.3%	50 52.1%	25 40.3%	44 58.7%
Non-faculty	100 37.6%	51 42.5%	49 42.6%	22 36.1%	53 50%	25 37.3%	62 38.8%	31 51.7%	41 42.7%	34 54.8%	24 32%
Team	9 3.4%	6 5%	5 4.4%	1 1.6%	6 5.7%	4 6%	7 4.4%	3 5%	3 3.1%	2 3.2%	6 8%
Other	6 2.3%	1 .8%	3 2.6%	2 3.3%	1 .9%	1 1.5%	1 .6%	3 5%	2 2.1%	1 1.6%	1 1.3%
				Non-faculty / Team / Other Coordinators							
	115 48.9%	58 48.3%	57 49.6%	25 41%	60 56.6%	30 44.8%	70 43.8%	37 61.7%	46 47.9%	37 59.7%	31 41.3%
# Responses	235 100%	120 100%	115 100%	61 100%	106 100%	67 100%	160 100%	60 100%	96 100%	62 100%	75 100%

Table 4
Workload Arrangements by Institution and Unit Characteristics

Workload Arrangements	All	Public/Private (223 responses)		Carnegie (222 responses)			Unit Accreditation (209 responses)		# Graduates (221 responses)		
		Public	Private	Bach	Mast	Doc	Yes	No	< < 100	100–249	250 or more
Fulltime duties	42 15.8%	22 19.3%	20 18.3%	6 10.5%	23 22.3%	13 21%	28 18.3%	14 25%	16 17.6%	14 23.7%	12 16.9%
Part-time (faculty–release time)	62 23.3%	37 32.5%	25 22.9%	18 31.6%	25 24.3%	18 29%	44 28.8%	13 23.2%	18 19.8%	17 28.8%	27 38%
Extra compensation	6 2.3%	3 2.6%	3 2.8%	2 3.5%	2 1.9%	2 3.2%	4 2.6%	1 1.8%	3 3.3%	3 5.1%	
Part-time (job responsibilities)	99 37.2%	47 41.2%	52 47.7%	28 49.1%	45 43.7%	26 41.9%	68 44.4%	24 49.2%	48 52.7%	22 37.3%	27 38%
Overload / added responsibility	8 3%	1 .9%	7 6.4%	2 3.5%	6 5.8%		5 3.3%	3 5.4%	5 5.5%	1 1.7%	2 2.8%
Other arrangements	6 2.3%	4 3.5%	2 1.8%	1 1.8%	2 1.9%	3 4.8%	4 2.6%	1 1.8%	1 1.1%	2 3.4%	3 4.2%
# Responses	223 100%	114 100%	109 100%	57 100%	103 100%	62 100%	153 100%	56 100%	91 100%	59 100%	71 100%

three (1.1%) units have three to five and six or more full-time support persons respectively. This trend is prevalent across all institutional characteristics. Once again, these results indicate that education units tend to rely more on part-time personnel to support the assessment function. Forty-five (16.9%) report having three to five persons, and 70 (26.2%) having six or more persons.

What is the estimated annual cost of the supporting the assessment function?

Table 6 and Figure 1 present summaries of the estimated annual spending on assessment support by institutional and unit characteristics. As indicated in Table 7, the estimated annual spending of the 232 reporting institutions ranges from approximately $11.5 million to approximately $17.1

Table 5
Support Infrastructure – Fulltime and Part-time Personnel

		Public/Private		Carnegie			Unit Accreditation		# Graduates		
	All	Public	Private	Bach	Mast	Doc	Yes	No	< < 100	100–249	250 or more
				Fulltime							
		(223 responses)		(214 responses)			(202 responses)		(212 responses)		
< 3	215 80.8%	107 90.7%	108 93.9%	56 91.8%	100 95.2%	58 87.9%	146 92.4%	56 93.3%	91 94.8%	57 93.4%	64 87.7%
3-5	15 5.6%	11 9.3%	4 3.5%	4 6.6%	3 2.9%	8 12.1%	11 7%	2 3.3%	3 3.1%	3 4.9%	9 12.3%
6 or more	3 1.1%		3 2.6%	1 1.6%	2 1.9%		1 .6%	2 3.3%	2 2.1%	1 1.6%	
sub-total (Fulltime)	233 100%	118 100%	115 100%	61 100%	105 100%	66 100%	158 100%	60 100%	96 100%	61 100%	73 100%
				Part-time							
		(230 responses)		(229 responses)			(215 responses)		(227 responses)		
< 3	115 43.2%	57 49.6%	58 50.4%	23 37.7%	54 51.9%	37 57.8%	73 47.1%	34 56.7%	53 55.2%	23 38.3%	38 53.5%
3-5	45 16.9%	23 20%	22 19.1%	12 19.7%	15 14.4%	18 28.1%	33 21.3%	8 13.3%	15 15.6%	9 15%	19 26.8%
6 or more	70 26.2%	35 30.4%	35 30.4%	26 42.6%	35 33.7%	9 14.1%	49 31.6%	18 30%	28 29.2%	28 29.2	13 19.7%
sub-total (Part-time)	230 100%	115 100%	115 100%	61 100%	104 100%	64 100%	155 100%	60 100%	96 100%	60 100%	71 100%

Table 6
Summary of Annual Expenditure on Assessment Infrastructure

Expenditure	Frequency	Percent	Lower Cost ($000)	Upper Cost ($000)
< $25,000	84	31.6	[a] 1,050	2,100
$25,000 - $49,999	47	17.7	1,175	2,350
$50,000 - $74,999	35	13.2	1,750	2,625
$75,000 - $99,999	25	9.4	1,875	2,500
$100,000 - $149,999	22	8.3	2,200	3,300
$150,000 - $199,999	11	4.1	1,650	2,200
$200,000 - $249,999	5	1.9	1,000	1,250
$250,000 or more	3	1.1	750	[b] 750
	232	100.0	11,450	17,075

[a] conservative lower limit estimate based on half of $25,000 or $12,500
[b] conservative upper limit estimate using $250,000

million. As expected, public institutions in this sample tend to have larger programs, and also report higher annual costs than private institutions, with 37 (31.2%) spending at least $100,000 annually versus four (3.5%). Regardless of expenditure level, SCDE given resource constraints, and their inability to depend on the larger institutional reporting systems, are responding to assessment and accountability data management requirements.

Limitations and Conclusions

In addition to sample size, other study limitations related to the selection of institutions impact the extent to which these findings can be generalized. The sample fell short of the desired number of responses. Also, as previously mentioned, the sample size is 13 shy of the 279 responses required for a 95% confidence level. This issue is further compounded by missing data, which resulted in a smaller number of responses for some questions. The potential impact of additional re-dropping is unknown. However, a decision had to be made to complete the data collection cycle, which was just over one year. Further, the list of targeted institutions was compiled from the US News and World report database. The comprehensiveness and accuracy of this database is unknown. Finally, institutions in the second and third administration had the opportunity to respond to an improved but slightly longer (six questions) version of the survey. The more significant of the two issues is the addition of the response option "ability to aggregate and disaggregate data." This option was missing from the first version due to researcher oversight. Knowing that this is a critical data analysis requirement of assessment systems, the decision was made to add it to improve the survey. Based on the responses, it is reasonable to believe that more respondents on the first administration would have made this choice if it was available. Therefore, there is strong indication that the importance and relevance

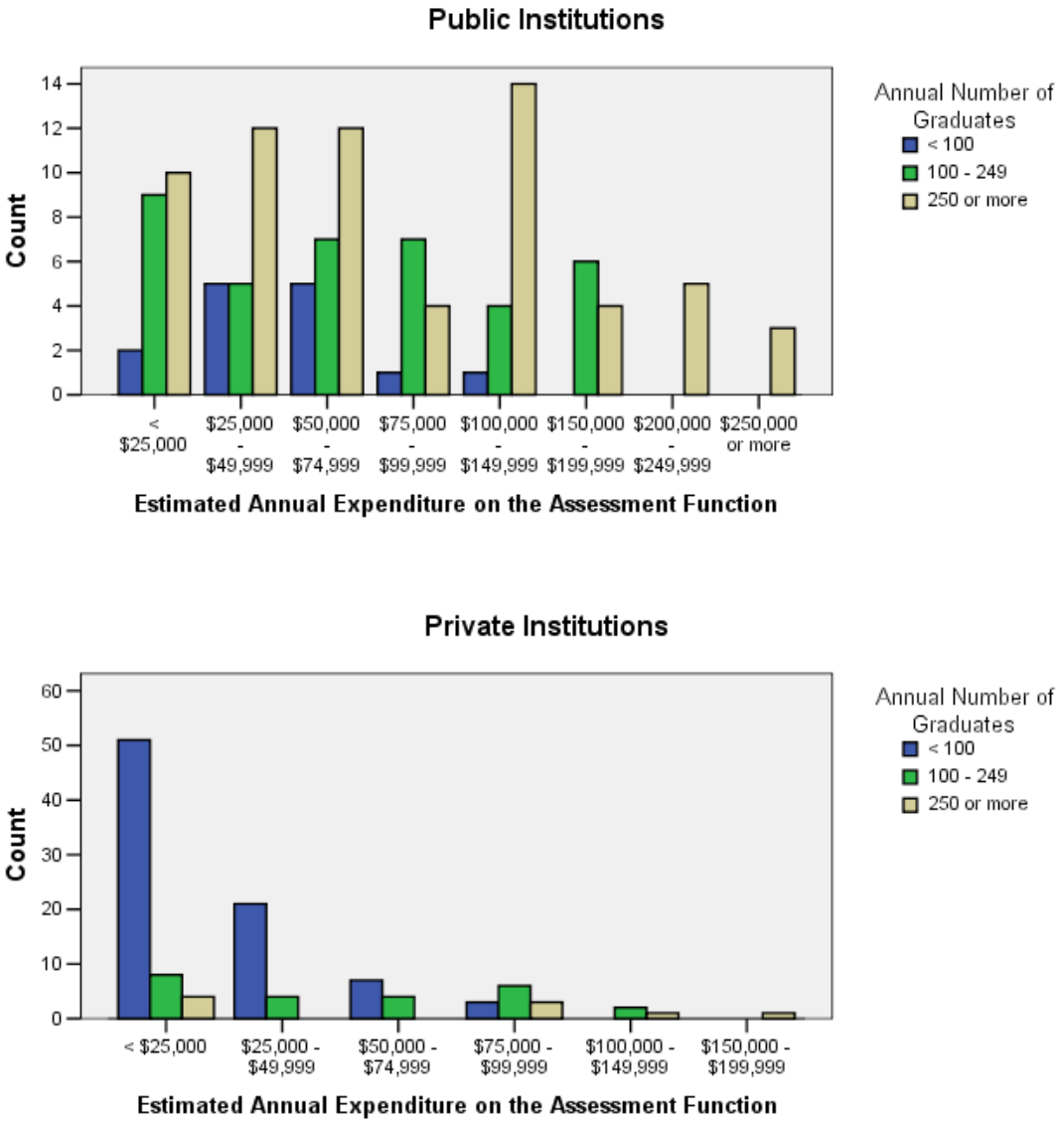

Figure 1. Estimated annual assessment costs by Institution Characteristics

of this reason was under-reported. The latter issue—the additional questions—was mitigated by the fact that pages one to four were the subject of the analysis and the three additional questions (outlined in the methods section) were not included in this analysis. However, it is also difficult to determine the full impact of the modified form in terms of its length. Despite these limitations, some insight may be gleaned from this initial report.

None of the responding units relied exclusively on the institutional level data management system, indicating that university-level systems are not capable of responding to all accountability and assessment requirements. All developed or are in the process of developing systems to satisfy accreditation and accountability data management. These systems ranged in complexity from paper-based to propriety systems developed by a few SCDE. Also, the results indicate that a significant number of SCDE are increasingly relying on third-party assessment software products to support their assessment endeavors. Relatively few—mostly private, unaccredited, smaller units—are relying on primarily paper-based systems. The implementation of these systems, including the popularity of third-party packages, further support the notion that university-level systems are inadequate and thus educational units are not benefiting from institutional infrastructure. Many SCDE indicated that they made choices that allowed them to respond to assessment and accountability requirement given respective resource constraints. The three primary reasons that account for assessment system choice are ease of use, cost, and the ability to customize. Also, the ability to aggregate and disaggregate data is important to many units. It is evident that the ability to conduct multiple analyses on data is a priority for many educational units as they seek to demonstrate their candidates are meeting expected outcomes. These reasons are additional evidence that existing university-level IT infrastructure does not adequately support data management and reporting that accompanies accountability mandates required of SCDE.

Faculty at more than half of the responding institutions manages the assessment function. More than one third of "faculty assessment coordinators" are untenured. Faculty work, as valued within the tenure process, may not align with assessment and accreditation work (Keil & Haughton, 2007). Therefore, it is difficult to ascertain the extent to which untenured faculty's involvement in assessments will be viewed as scholarly activity. It is also unclear how untenured faculty assessment coordinators will be able to navigate and manage these competing and potentially unrelated interests.

Unlike the centralized IT function in higher education, the assessment coordinator role is largely part-time. As with the coordinator's role, SCDE are heavily dependent on part-time support personnel. It is unknown whether the assessment system and related functions are meeting their optimal performance. It is our belief that given the inadequate staffing levels, units are not in a position to make this determination. In addition to resource constraints, this reliance on part-time support may also be a result of institution culture, the tendency to delay focus until an accreditation visit is on the horizon, or some other unknown factor.

Despite resource constraints, many SCDE are making relatively large investments in assessment systems with many spending at least $100,000 annually. As expected, public, doctoral granting, NCATE accredited units with large numbers of graduates tend to spend more. The reported costs were estimated and as such may not represent the true costs (e.g., opportunity costs) related to this expenditure. Also, the impact of money spent on the assessment function in terms of system effectiveness and unit quality outcomes is yet to be determined.

So what does all of this mean? SCDE, by and large, are responding to the accountability mandates, and are making varying levels of investments in their own unit level assessment systems and infrastructure. While the degree and specific type of institutional support is unclear, the evidence indicates that many units are shouldering a significant amount of the accountability mandates and the accompanying burdens on their own. This has resulted in a variety of assessment systems. Some may argue that the least sophisticated, paper-based system potentially bare little resemblance to the most sophisticated proprietary systems—though both were designed to serve the same purpose. Additional and more contextually based studies must be conducted to determine the level and type of institutional support as well as other hidden costs related to these accountability systems and infrastructure. Most of all, studies must be conducted to determine the effectiveness and the level of impact of these systems on quality outcomes. Of critical importance is the extent to which these data-driven changes and outcome assessments are impacting the continuous improvement requirement of education units and the relationship between these SCDE-level and institutional-level quality indicators.

Responding to the accreditation and accountability mandate to invest in data management systems and infrastructure that facilitates the collection, analysis, and reporting of data may be necessary, but not sufficient, for the preparation of highly qualified educators. This goal becomes even more complex and illusive given the lack of agreement about what constitutes a qualified educator. Berliner (2005) states, "Under the best of circumstances, it would be difficult to define a quality teacher; under political mandate to do so, it is likely to lead to silly and costly compliance-oriented actions by each of the states" (pp. 206-207). By default units of education are also likely to engage, and may have already engaged, in similar "silly and costly compliance-oriented actions."

References

Allen, I. E., & Seaman, J. (2007). Changing the landscape: More institutions pursue online offerings. *On The Horizon*, *15*, 130-138.

Alreck, P. L. & Settle, R. B. (2004*). The survey research handbook* (3rd ed.). New York: Irwin/McGraw-Hill.

Bradburn, E., & Zimbler, L (2002). *Distance education instruction by postsecondary faculty and staff: Fall 1998.* (Report No. NCES 2002-155). Retrieved on August 16, 2008, from National Center For Education Statistics: http://nces.ed.gov/pubs2002/2002155.pdf

Berliner, D. (2005). The near impossibility of testing for teacher quality. *Journal of Teacher Education*,*56*, 205-213.

Carey, K. (2007). Truth without action: The myth of higher-education accountability. *Change*, *September/October*, 24-29.

Diez, M. E. (2001). Assessing student competence in teacher education programs. In C. A. Palomba & T. Banta (Eds.), *Assessing student competence in accredited disciplines: Pioneering approaches to assessment in higher education* (pp. 29-48). Sterling, VA: Stylus Publishing.

Hawkins, B., & Rudy, J. (2007). *Educause core data fiscal year 2006 summary report.* Educause. Retrieved August 28, 2008, from http://net.educause.edu/ir/library/pdf/pub8004.pdf

Keil, V., & Haughton, N. (2007). Accreditation data collection requirements versus faculty loads: One college's use of self-study to balance these two realities. *The Teacher Educator*, *42*, 209-223.

National Center for Education Statistics. (n.d.). *Fast facts.* Retrieved on August 27, 2008, from National Center For Education Statistics: http://nces.ed.gov/fastFacts/display.asp?id=98

National Council for the Accreditation of Teacher Education (2008). *Professional standards for the accreditation of schools, colleges, and departments of education.* Washington, DC: National Council for Accreditation of Teacher Education.

Palomba, C. & Banta, T. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Francisco: Jossey-Bass.

Teacher Education Accreditation Council (2006). TEAC accreditation process. Retrieved September 29, 2006, from http://teac.org/accreditation/summaryoutline.asp

Noela A. Haughton is an Assistant Professor at the University of Toledo. Her research interests include assessment and evaluation in higher education, including instructional and educational systems. E-mail: nhaught@utoledo.edu.

Virginia L. Keil is an Associate Professor at the University of Toledo, and Director of Teacher Education. Her research interests include assessment and evaluation in teacher education, effective models of clinical supervision, and school leadership. E-mail: Virginia.keil@utoledo.edu.