Journal of Research in Education
Volume 24, Number 1
Spring/Summer 2014

A Journal Published by the
Eastern Educational Research Association

Table of Contents

Past Editors of the Journal

2013-2014	Andrew Shim University of South Dakota
2012-2013	Abbott Packard University of West Georgia
2008-2012	Michael Miller University of Arkansas
2004-2008	Darla Twale University of Dayton
1996-2004	Donald Biggs SUNY, University at Albany
1990-1996	Betty Yarborough Old Dominion University

From the Editor

I am pleased to present the most recent volume and issue of the Journal of Research in Education. So much of the good work of the Eastern Educational Research Association, including this journal, comes from individuals who give selflessly of their time and energy. This includes serving as reviewers and editors of the journal, and I must thank our past two editors for the time they gave to help the journal be successful. Dr. Andy Shim and Dr. Abbot Packard both game a tremendous amount of time and energy to the journal, and we are all grateful for their commitment.

I am also pleased to announce that beginning January 1, 2015, Dr. Barbara Kuwulich and Dr. Mary Alice Varga, both from the University of West Georgia, will become editors for the Journal. Their combined expertise, attention to detail, and commitment to EERA are much appreciated and we all look forward to their work during the next four years.

There are also many individuals who I must thank for their reviewing of manuscripts. They are listed in the following pages. For this issue of the journal, we had an 18% acceptance rate, and I think that you will find the articles contained in the journal to be an impressive representation of the academic work of these authors.

I encourage you to continue to consider the Journal of Research in Education as an outlet for your academic work, and I also hope that you will consider participating in the upcoming EERA conference in Sarasota, Florida this February.

Best wishes for the fall term,

Michael T. Miller
Interim Editor

Journal of Research in Education
2014-2015 Reviewers

Mary Alice Barksdale
Virginia Tech University

Jan Blake
University of South Florida

Kenda S. Grover
University of Arkansas

Jeanne Jenkins
John Carroll University

Dennis King
University of Maine

Hsiu-Lien Lu
Georgia Southern University

Adam A. Morris
Crowder College

Stephen Nelson
Bridgewater State University

Richard E. Newman
Presbyterian College

Abbot Packard
University of West Georgia

John Queener
University of Akron

Susan Ramlo
University of Akron

David Shannon
Auburn University

Christine Shea
East Carolina University

Melanie Shores
University of Alabama, Birmingham

Everrett A. Smith
University of Arkansas

Teacher Empathy and Its Relationship to the Standardized Test Scores of Diverse Secondary English Students

Timothy B. Bostic
Culinary Institute of Virginia

Correspondence related to this article should be directed to Tim Bostic at tbsotic@chefva.com

Abstract

The purpose of this research study was to ascertain whether there is a relationship between teachers' cognitive role taking aspect of empathy and the Virginia Standards of Learning (VSOL), English/Reading scores of their students. A correlational research design using hierarchical multiple regression was used to look for this relationship. In order to control for variables previous research has shown to contribute to student achievement, a teacher's years of experience, degree level, self-efficacy beliefs about managing classroom behavior and a teacher's expectations for her students were measured and placed into the regression equation. The study attempted to see if the relationship was stronger based on the ethnicity and course level of the students. The results indicate that the null hypothesis cannot be rejected. However, the results also indicated that the other teacher variables for which this study controlled were also not contributing to the variance in the test scores. These findings led to the conclusion that standardized tests, by their very nature, may possibly not be susceptible to teacher attributes or dispositions. Further, it was concluded that teachers may need to acknowledge that the VSOL, English/Reading tests may be measuring a very small part of student achievement which, in many cases, can be considered learning to pass the tests.

Currently, research has become more complex as researchers try to identify the specific personal characteristics and teacher dispositions that help teachers facilitate successful learning in their classrooms. Rice (2003), in her meta-analysis of research on what teacher attributes make teachers effective, found significant gaps in knowledge about which teacher characteristics help to make them most effective, especially at the secondary level. However, one characteristic that proves beneficial to student learning is teachers' empathy. For the purposes of this study empathy is operationalized as understanding a student's perspective. It is assumed that when teachers understand their students, they will care about them, their cognitive and affective growth and well being. Additionally, it is assumed that this understanding will make teachers more effective in terms of helping students achieve academically as measured by standardized tests. Further, for students who have a history of academic failure, teacher empathy and caring is even more important in terms of motivating them and connecting them to academic opportunities.

The purpose of the study was to measure teachers' empathy and to assess it contributions to student performance on the English reading scores on the Virginia Standards of Learning test (VSOL) for eighth and eleventh grade students. The study examined the effects of teacher empathy on academic performance for different student groups. The study also examined

classroom management skills because it was suspected that they might interact with teachers' expectations for their students and empathy for them.

Significance of the research

Sociological research has found that the greater the social difference between teachers and their students, the lower the teacher's expectations will be for those students, and they will be less supportive and positive in their interactions with students they perceive to be of lower social status than themselves (Metz, 1990). Recent research has begun to examine how empathy makes teachers more effective working with diverse student populations (Darling-Hammond, 2000; McAllister and Irvine, 2002). Studies by Noblit, Roger and McCadden (1995) and Noddings (1995) indicate that many students must believe that their teachers care about them and understand them in order for them to be motivated to learn. Teacher empathy is thought to affect students indirectly by altering the learning environment. Empathetic teachers promote a caring climate in the classroom that makes students feel more connected to the school. Also, teachers who understand their students are more likely to be motivated to make lessons relevant to their students, and this then increases students' interest and engagement in learning.

In recent decades, the student population in U.S. schools has increased in diversity. In 2000, 65 percent of the K-12 student population was non-Hispanic white and 35 percent were Hispanic, Black, Asian/Pacific Islander and Native American (Fowler, 2004). While the diversity of the student population has been increasing, the diversity of the teaching staff of elementary and secondary schools has been decreasing. Twelve percent of the graduates of teacher preparation (education) programs in 1977 were African American and by 1993 the proportion had decreased to 10 percent (Ohio State Office of Educational Oversight, 1997). In 2000, 86 percent of all elementary and secondary teachers were white (Gay and Howard, 2000). Given the increasing social and cultural disparity between teachers and students, the need to understand how teacher empathy affects instructional practices and the learning environment is critical. If white middle-class teachers, which comprise the majority of the teaching population, are less empathetic toward ethnic minority students (due to social and cultural factors), then these students will be disadvantaged by exposure to a learning environment that is less supportive of them than that experienced by their white peers. Additional research on teacher empathy is needed to prevent social/cultural disparities between teachers and students from resulting in disparities in instructional support and learning opportunities.

Review of Literature

Considerable educational research has examined how different teacher characteristics affect learning, including teacher empathy. The most current research is reviewed in this section. A discussion of how empathy is defined and how specific constructs of empathy can be measured also is included.

Teacher Characteristics and the Learning Environment

Prior research indicates that teachers who have mastered their subject matter understand their subject matter and have sound pedagogical training are more successful in helping their students learn. Darling-Hammond (2000) found that teachers who have a thorough understanding of learning theories, teaching methods and curriculum knowledge are more effective in promoting student learning in diverse classrooms, that is "teaching from the perspective of learners who bring diverse experiences and frames of reference to the classroom" (p. 166). There is more to effective teaching, however, than just understanding course content and how to present it. Teachers need to be able to create learning environments that are conducive to learning.

How teachers express caring and concern for their students is not important. What is important is that they let their students know they care (Patterson & Purkey, 1993). Pierce (1994) found that teachers who showed respect, caring and closeness to students at risk lessened the chance that they would fail. By interacting with students in a caring and respectful manner, teachers can motivate students to care about academic success and foster the belief that they can be successful in school. Teachers who enter the profession because they have a desire to make a difference in the lives of their students and want them to succeed academically have better student outcomes (in terms of grades and standardized test scores) than teachers who are motivated by other factors (Gordon, 1999; Kuklinski & Weinstein, 2001).

Empathy and Student Learning

The effect of teacher empathy on learning has been established; teachers who students perceive as empathetic have higher learning outcomes (McAllister & Irvine, 2002). When students feel that their teachers are trying to understand them as individuals and are concerned about them and their life, they work harder and achieve more academically (Coffman, 1981). McAllister and Irvine (2002) found when teachers show tolerance and acceptance of their students, they create a better learning environment. This finding is supported by research that resulted in the Wingspread Declaration, which promotes caring environments in schools (Viadero, 2004). This prior research indicates that a caring school environment not only affects academic performance but also reduces anti-social behavior (Viadero, 2004). Apparently, students learn to treat others in the manner they have been treated.

Teaching can be viewed as a series of social interactions between students and teachers. Because the teacher is the authority figure in the room, the teacher, whether consciously or unconsciously, teaches by example. It is imperative, therefore, that the teacher model for students appropriate ways of interacting with other people. By illustrating empathy, the teacher establishes a positive climate of social interactions in the classroom that are conducive to learning while at the same time modeling positive social interactions for students that reduce anti-social behaviors that disrupt the learning environment (Sandven, 1979). Empathy, then, has an impact on learning both cognitively and affectively. By modeling how to work effectively with different types of people, teachers help students acquire the ability to understand others' perspectives when the situation requires it and promote cross-cultural understanding. Students are better able to establish a connection with caring teachers, and to internalize their teachers' values. This then may motivate students to produce better work for their teachers.

Defining Empathy

Demos (1984) defines empathy as an affective psychological stance towards another person. People often react emotionally to the perceived distress of others, but researchers differentiate between empathy, which is an affective psychological stance, and sympathy or pity (Davis, 1990; Batson, Fultz & Schoenrade, 1987). Empathy for the purposes of research, including this study, is defined as the ability to understand another person's perspective. This definition has been used by Everding and Huffaker (1998), Davis (1990), Reed (1984), Ickes (1997), Lanning (1991), as well as Stotland, Mathews, Sherman, Hanson and Richardson (1978). This specificity is in keeping with experts in the field noting the importance of having a clear definition of empathy before attempting to measure it (Feldstein & Gladstein, 1980). Empathy, as a psychological construct, is illustrated in the following two vignettes.

Vignette 1: An Empathetic Teacher

> Eric, a student from an economically depressed neighborhood, has been tardy for four days in a row. When his teacher asks him why has been late he tells his teacher he walks his sister to her middle school first, in order to protect her from local gangs that have threatened her. He gets her to school twenty minutes early, but he gets to school twenty minutes late. The teacher explains to Eric that she can appreciate what he is going through, but he is missing out on important instruction every day. She asks his guidance counselor to alter his schedule so that he has a study hall during the first bell of the day and has him moved to a later section of her English class. In this instance, the teacher was willing to find out the cause of Eric's problem and help find a solution that enabled him to succeed academically.

Vignette 2: An Un-empathetic Teacher

> Janice has been more than twenty minutes late almost every day for her English class. The teacher warns her that after twenty minutes she is considered absent even if she shows up, and after ten absences in one semester, she automatically fails the class. When Janice tries to explain, the teacher tells her the explanation is unimportant; she must get to class on time. By the end of the semester Janice has been late more than ten times and her teacher fails her. When Janice appeals the teacher's decision, the appeals committee learns that her mother is a drug addict and that Janice cannot sleep at night because of people constantly entering and exiting the house. The committee excuses her absences, and Janice remains in her English class. Because Janice's teacher was unwilling to understand her perspective and to find out whether she had a valid reason for being late to class, Janice becomes angry and does not put forth much effort during the second semester. Thus, the teacher's lack of empathy hindered Janice's motivation and academic performance in her English class.

Measuring Empathy

As with any psychological construct, obtaining a valid and reliable measure of empathy is a difficult task. Macarov (1978) describes initial efforts to measure it. The first attempt to measure

empathy was in 1949 by Dymond and continued until Truax's effort in 1961. In their review of several different measures, Feldstein and Gladstein (1980) found that none of the instruments were valid and reliable measures of empathy (i.e., "internal and unobservable activation of the counselor's feelings and fantasies") or the communication of the empathic experience (i.e., "checking and altering their [the counselor's] statements to maintain an accurate understanding of the clients"). However, some of the instruments were able to measure other aspects of empathy, such as, cognitive empathy, defined as role-taking skills, which is the aspect of empathy pertinent to this study.

Davis (1979) developed the Interpersonal Reactivity Index (IRI), a 28-item Likert scale that contains sub-scales that measure four aspects of empathy. The aspect most relevant to this study is the Perspective-Taking sub-scale. The IRI has reasonably good psychometric properties: the internal reliabilities range from .71 to .77; the test-retest reliabilities range from .62 to .80 (Davis, 1979). The Perspective Taking sub-scale of the IRI has a Cronbach's alpha of .64 (Constantine and Gainor, 2001). The properties of the IRI, while low, were acceptable for this study, and the Perspective taking sub-scale matched the study's operational definition of empathy.

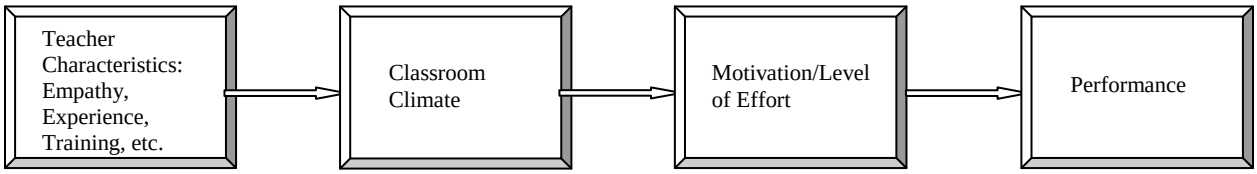
Design and Methods

A non-experimental, correlational design was used to examine the relationship between teacher empathy and students' academic performance as measured by the VSOL's, Virginia's standardized achievement test (McMillan, 2000; Agresti & Finlay, 1997). The study addressed the following questions:

1. What is the effect of teacher empathy on students' academic performance?
2. Is the effect of teacher empathy on student academic performance consistent across all ethnic groups?
3. Is the effect of teacher empathy on student academic performance consistent across academic ability groups?

The underlying logic for these questions is represented in figure 1.

Figure 1. Logic Graphic



Instrumentation

Measures of teacher characteristics

The instruments used to measure teacher characteristics included a self-administered survey to collect background information and several scales developed for the study to measure the teacher attributes of interest. Several teacher background characteristics were measured because they

were expected to affect student academic performance, and therefore, their effect on this outcome variable had to be controlled for in the analysis. These predictor variables included: (a) teaching experience; (b) professional training (i.e., method of certification and graduate training); (c) classroom management ability; and (d) expectations for students. Teachers were surveyed to collect information about their educational background, teacher training, professional credentials, and teaching experience.

Three scales were developed to measure teacher attributes: (a) teacher empathy; (b) teacher self-assessment of classroom management ability; and (c) teacher expectations for students. The scale used to measure empathy, which was operationally defined as cognitive role-taking (Feldstein & Gladstein, 1980), was an adaptation of a sub-scale from the Interpersonal Reactivity Index developed by Davis (1979). The Cronbach alphas for the adapted scale ranged from .67 to .76. The Teacher Interpersonal Self-Efficacy Scale developed by Emmer & Hickman (1991) was adapted to measure teachers' perceived self-efficacy in managing students' classroom behavior. The adapted scale had a Cronbach alpha of .92. A scale to measure teachers' beliefs about expectations for students' ability to learn was based on research by Gottfredson, Marciniak, Birdseye and Gottfredson (1995), Harris and Rosenthal (1985), and Jussim and Eccles (1992). The adapted scales were field tested with a sample of in-service teachers and the reliability coefficients were comparable to those reported in previous studies.

Measures of student performance

The Virginia Standards of Learning (VSOL) in the areas of English, mathematics, history, social science, and science are intended "to set reasonable targets and expectations for what teachers are expected to teach and students are expected to learn" (see the Virginia SOL Technical Manual, May 2000, page 1). The purposes of the educational assessments at selected grades 3, 5, and 8) and high school subjects are to inform parents and teachers about what students are learning in relation to the VSOL and to hold schools accountable for teaching the VSOL content. For this study, the English, Reading test scores, a criterion-referenced test aimed at assessing the efficacy of instruction on state mandated learning objectives in use word analysis strategies and information resources and demonstrate comprehension of printed material, were used. The eighth grade test consists of forty-five multiple choice items and the eleventh grade test consists of fifty multiple choice items (http://www.pen.k12.va.us/VDOE/Assessment). However, the actual number of questions may vary on each administration of the test due to the need to field test some items. The test, from which data for this study were obtained, was administered in the Spring of 2005. Equating was done to ensure that all forms of the test were of equal difficulty. Every time a new test form is constructed, attempts are made to make the new form equal in difficulty to previous forms. This process was accomplished through data collected during field tests. The data collection design was Design IV procedure for common item, non-equivalent groups (Angoff, 1971). To explain this scoring procedure further, item parameters, developed using Rasch equating procedures, for all forms were on the same Rasch ability scale. People within a certain range of ability should have the same ease or difficulty in answering questions on administrations of different tests. "The parameter estimates for each form were placed on a

common metric by using the Rasch equating constant procedure. This resulted in the item parameters for *all* forms being on the same Rasch ability scale."[1]

Sample

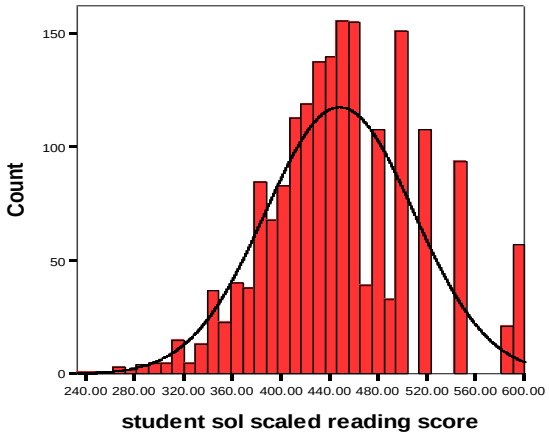
The sample of teachers was drawn from five high schools and eight middle schools located in an urban school district in southeastern Virginia. Twenty-seven of the 178 eighth grade and eleventh grade English teachers volunteered to participate in the study. The study sampled eighth grade and eleventh grade teachers because Virginia assesses student performance in English/Reading at these grade levels. Because all students are required to take the state's standardized achievement test, it was selected as the measure of student performance. All of the teachers in the sample were licensed and certified to teach English in the Commonwealth of Virginia. Only four of the teachers had received their teacher licensure through alternative means. Thirty percent of the teachers had earned a Master's degree in education or in English. On average, the teachers had a little over twelve years of teaching experience.

The sample of students included 1,861 students who were enrolled in eighth grade and eleventh grade in the spring of 2005. The sample consisted of 853 males (46%) and 1,008 females (54%), and was ethnically diverse. Slightly more than half (55%) were African American, approximately a third were Caucasian (35%), and a small proportion were Latino (3%), Asian-American (3%) or classified as another ethnic group (4%). Seventy-nine percent of the students in the sample were eighth graders (1,475) and 21 percent (386) were eleventh graders. Sixty-five percent (1,216) of the sample was enrolled in regular English classes, 30 percent (566) was enrolled in honors English, and a small percentage were taking remedial English (79, 4%).

The sample population of students in this study had an average SOL; English, Reading score of 447.35. The minimum score was 233 and the maximum score was 600. The passing score for this measure is 400; however, students who score above 500 are considered to have "passed advance." The standard deviation was 60.88. The scores are skewed as illustrated by Figure 2.

[1] For a complete discussion of how the Virginia Standards of Learning are scored, administered and equated please refer to the Virginia Standards of Learning Assessments: Technical Report available at www.pen.k12.va.us/VDOE/Assessment.

Figure 2. Histogram of Students' SOL; Reading, English Scores



For the eighth grade students in the sample population seventy-eight percent (78%) passed the VSOL English/Reading test. For the eleventh grade students in the sample population ninety-four percent (94%) passed the VSOL English/Reading test. For the district as a whole, the eighth grade students had a pass rate of seventy-four percent (74%). The eleventh grade students had a pass rate of ninety percent (90%). Thus, these differences indicate that the sample population outperformed the total population in both the eleventh and eighth grades.

Analysis

A hierarchical multiple regression analysis was performed on the predictor variables and the outcome variables and VSOL English/Reading test to test the relationship between the. Hierarchical regression was used for the analysis because the theory linking the variables of interest requires that variables be placed in the regression equation in a specific order rather than allowing the statistical program to determine the order. The regression equation then had the Teacher Interpersonal Self-Efficacy data added to see if more variance could be accounted for, and the data on teacher expectations for students was added. A regression analysis was conducted to examine possible sub-group differences for students based on their ethnicity and their academic level placement (i.e. regular English, honors, remedial). to look for a relationship between the VSOL, English/Reading scores and the teachers' empathy.

Results and Discussion

The analysis of the teacher data indicate that a teacher's empathy is positively correlated ($p <$ .01) to a teacher's years of teaching experience and a teacher's expectation for students. A teacher's empathy is negatively correlated ($p < .01$) to the type of degree they have (i.e. bachelors, masters, etc.), certification and their self-efficacy beliefs about ability to manage classroom behavior. A teacher's belief in her ability to manage classroom behavior is positively correlated ($p < .01$) to years of teaching experience and expectations for students. A teacher's belief in his ability to manage classroom behavior is negatively correlated ($p < .01$) to the type of degree, being regularly certified and teacher empathy. A teacher's expectation for his students is positively correlated ($p < .01$) to years of teaching experience, their type of degree, being

regularly certified, teacher empathy and a teacher's belief in his ability to manage classroom behavior. A teacher's expectation for his students is negatively correlated ($p < .01$) to the teacher's ethnicity. For the twenty-six teachers in this study, their empathy was positively correlated, although it was a low correlation, with their years of teaching experience, thus the more years of teaching the higher the empathy level. While a few of the other characteristics were significantly correlated statistically with empathy, none was above .30. This finding is important because it helps to illustrate that the longer teachers teach, the better able they are to understand their students' perspectives.

Table 1 provides the ANOVA for Teacher Empathy and Course Type. The data indicate a statistically significant difference ($p < .01$) between the three levels of courses offered, remedial, regular and honors. A Bonferroni Post Hoc test was performed which indicated that the mean difference between the remedial teachers' empathy and the regular and honors teachers' empathy was significant ($p < .01$). The mean difference between the regular and honors teachers' empathy was not significant.

Table 1

Means, Standard Deviations and One-Way Analyses of Variance for Teacher Empathy on Course Type with Effect Size between Regular and Workshop and Honors and Workshop

Variable	Teacher Empathy		ANOVA	
	M	SD	F (2, 1858)	η^2
Workshop	4.12	.65	6.48**	
Regular	3.90	.54		.41
Honors	3.89	.52		.44

Note: **$p < .01$

The effect sizes for these differences, while not large, was .41 for teachers teaching regular sections of English, compared to those working with workshop students and .44 for teachers working with honor students, again compared to the workshop teachers. The moderate effect size does indicate that there is a practical difference between the empathy levels for the teachers working with the different type of students. It is an encouraging sign that teachers who are working with the students who struggle the most possess a higher mean level of empathy. It is possible that because these teachers are able to indicate to their students that they understand their perspectives, they are able to help almost half (39 out of 79) of their students pass the VSOL, English/Reading test.

Effect of Teacher Empathy

In terms of the first and second question, what is the effect of teacher empathy on students' academic performance and is the effect on student academic performance consistent across all ethnic groups, the following results were found. The type of degree a teacher holds contributes negatively to the variance in test scores as indicated by the negative beta coefficient (-3.14; $p < .01$). The data indicate that the teacher obtaining a higher degree negatively impacts a student's

score for all of the students. The data indicate that teacher empathy is statistically significant for Caucasian students although it contributes negatively to the variance in test scores (-16.61; $\underline{p} < .01$). Additionally, for Caucasian students the relationship between teacher empathy and the SOL scores contributes as much to the variance in SOL scores as the years of teaching experience, the degree type and whether the teacher received their certification through a regular certification program. Empathy is not statistically significant for African-American students in the sample, nor is it statistically significant for the students in the sample regardless of ethnicity.

The final question of interest to this study was whether the effect on teacher empathy on student academic performance was consistent across academic ability groups. Here the analysis indicates that, for the students enrolled in a regular section of English, years of teaching experience, degree type of the teacher and being regularly certified are statistically significant ($\underline{p} < .01$). The teachers' self-efficacy beliefs and their expectations for students are also statistically significant ($\underline{p} < .05$). For the students in the sample enrolled in an honors section of English, years of teaching experience, degree type of the teacher and being regularly certified are statistically significant ($\underline{p} < .05$). The teachers' expectations and teacher empathy are both statistically significant ($\underline{p} < .01$). Further, teacher empathy is making the largest contribution to the R^2 (.03; $\underline{p} < .01$) change. For students enrolled in remedial classes, none of the constructs are making statistically significant contributions to the R^2 change.

The analysis does not indicate that empathy is related to student performance. There are a number of reasons why no association was found. The study had a number of limitations. The first was the method and instrument used to measure empathy. Responses on self reported measures are affected by respondents' feelings about what is socially acceptable. This often affects the variance in responses. The amount of variance on the Empathy scale was small. Also, the student sample was fairly high performing since the vast majority of the sample population passed the test; thus, the amount of variance on the student data was small. Additionally, since the sample population out performed the population as a whole (70% vs. 74% for the 8th graders and 90% vs. 94% for 11th graders), it is possible that the high level of empathy for the teacher sample was a contributing factor. Finally, the measure of achievement used may not be influenced by the teacher variables that previous research studies indicate impact student achievement. By their nature standardized achievement tests may not be as sensitive a measure as needed for the purposes of this study. Other student outcomes, such as level of effort, might be better outcome measures.

When narrowly defining empathy, which is considered an aspect of caring, the small piece of it measured for this study did not make a difference in student achievement as measured by the VSOL, English/Reading test. Other constructs that have been shown repeatedly to make a difference in student achievement also are not making much of a contribution to the variance in the students' scores. It may be an issue centered on the student achievement measure. In an attempt to make school systems accountable for student learning, policy makers have demanded objective measures of achievement. Standardized achievement tests like the VSOL, English/Reading test may not be appropriate for assessing the impact of teacher characteristics on student learning because they are designed to measure instructional effectiveness and presentation of specific curriculum content. These measures may not be influenced by teacher characteristics because they provide results for a very narrow understanding of education. They

are more reflective of the training students receive in a specific content area rather than a deep appreciation for the content in terms of how it helps the students understand themselves and the world in which they live.

Conclusions

This study attempted to ascertain the relationship between teacher empathy and VSOL, English/Reading scores for eighth and eleventh grade students. It did not find a relationship between the standardized test scores and teacher empathy. Teacher characteristics which in past research studies have had a relationship to test scores, such as years of experience and self-efficacy beliefs about managing classroom behavior, also did not have a relationship to the VSOL, Reading/English scores. This fact created some consternation in terms of why this might be the case. While the possible threats to the measurement of teacher empathy are reasons for why the findings were non-supportive of the research hypotheses, the other measures were not subject to those same threats (i.e. years of experience, self-efficacy beliefs, degree attainment, etc.). The small teacher sample does not make it possible to conclude definitively that the VSOL, English/Reading test is or is not a measure that is influenced by these teacher traits.

Though the study's findings did not support the expected connection between teacher empathy and student achievement, the difference in the percentage of students who passed in the district and the percentage of students who passed in the sample population may be attributable to the high level of empathy of the teacher sample. The eighth grade African-American students in the district had a pass rate of seventy percent (70%), and the eighth grade African-American students in the sample population had a pass rate of seventy-nine percent (79%). The eleventh grade African-American students in the district had a pass rate of seventy-eight percent (78%), and the eleventh grade African-American students in the sample population had a pass rate of eighty-eight percent (88%). This finding may be practically significant for the district if they are concerned with finding ways to close the achievement gap between their Caucasian and African-American students. With the increasing disparity between the teacher population and the student population for the nation as a whole, teacher empathy needs to continue to be studied in terms of its ability to affect student achievement. Future studies should attempt to avoid the limitations of this study. Special attention should be placed on obtaining a more varied sample of teachers on the construct of interest. Additionally, it would be recommended that more than one measure of student achievement be used.

Empathy, as operationalized and measured for this study, did not contribute to the variance in the test scores of diverse students in practically significant ways. However, it is important to note that teachers who work with students in remedial sections of English eight and eleven had higher mean empathy scores than those working with students in honor and regular sections of English. This finding seems to illustrate the reciprocal nature of the relationship between students and teachers because teachers working with these types of students may be better able to understand their students' perspectives. As a classroom teacher and a researcher, I still believe that students need to believe their teachers care about them and their academic progress. The qualitative research and the anecdotal findings are too strong to indicate otherwise.

The fact that the research hypotheses were not proven should not dissuade researchers from continuing to explore and research the connection between caring and academic achievement. Noddings (1992, 1995, 2002) argues that even the meager success of academic achievement can not be attained if students do not feel cared for. We have a responsibility to ensure that all students in our nation's public school system receive the type of education that will prepare them to become who ever and what ever they so desire. Students are entering a world that is becoming extremely complex and competitive. It is not that the tests in themselves are problematic. Policy makers have the right to require that students can prove themselves trained to accomplish basic reading and writing tasks. However, when policy makers place so many negative consequences on schools and students who are unable to pass these tests, it promotes teachers and administrators to solely focus on helping the students pass these tests. For students who struggle to pass these tests, it is understandable that schools may inadvertently become so caught up in getting the students to pass that they lose sight of the overarching goals of education.

References

Angoff, W. H. (1971). *Educational Measurement 2nd edition.* Washington, DC: American Council on Education.

Agresti, A. & Finlay, B. (1997). *Statistical methods for the social sciences* (3rd ed.). Upper Saddle River: Prentice Hall.

Batson, C.D., Fultz, J., & Schoenrade, P.A. (1987). Adults' emotional reactions to the distress of others. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp.163-184). Cambridge: Cambridge.

Coffman, S. L. (1981). Empathy as a relevant instructor variable in the experiential classroom. *Group & Organization Studies*, *6*(1), 114-120.

Constantine, M. G. & Gainor, K. A. (2001). Emotional intelligence and empathy: Their relation to multi-cultural counseling knowledge and awareness. *Professional School Counseling*, *5*(2),131-137.

Darling-Hammond, L. (2000). How teacher education matters. *Journal of Teacher Education*, *51*(3).

Davis, C.M. (1990, November). What is empathy, and can empathy be taught? *Physical Therapy,* 70(11), 707.

Davis, M. H. (1979). Individual differences in empathy: A multidimensional approach (Doctoral dissertation, University of Texas, Austin, 1979). *Dissertation Abstracts International,* 40, 07B.

Demos, V. (1984). Empathy and affect: Reflections on infant experience. In J. Lichtenberg, M. Bornstein & D. Silver (Eds.), *Empathy II* (pp. 9-34). Hillsdale, NJ: Erlbaum.

Emmer, E. T., & Hickman, J. (1991). Teacher efficacy in classroom management and discipline.

Educational and Psychological Measurement, 51, 755-765.

Everding, H. E. & Huffaker, L. A. (1998). Educating adults for empathy: Implications of cognitive role-taking and identity formation. *Religious Education, 93*(4), 413-430.

Feldstein, J. C. & Gladstein, G. A. (1980). A comparison of the construct validities of four measures of empathy. *Measurement and Evaluation in Guidance, 13*(1), 49-57.

Fowler, F. C. (2004). *Policy studies for educational leaders: An introduction* (3rd ed.). Upper Saddle River, NJ: Pearson Education.

Gay, G. & Howard, T. C. (2000). Multicultural teacher education in the 21st century. *Teacher Educator, 36*(1), 1-16.

Gordon, G. L. (1999). Teacher talent and urban schools [Electronic version] *Phi Delta Kappan, 81*(4), 304.

Gottfredson, D. C., Marciniak, E. M., Birdseye, A. T., & Gottfredson, G. D. (1995). Increasing teacher expectations for student achievement. *The Journal of Educational Research, 88*(3), 155-163.

Harris, M. J. & Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: 31 meta analyses. *Psychological Bulletin, 97*(3) 363-386.

Ickes, W. (1997). Introduction. In W. Ickes (Ed.). *Empathic Accuracy*, (pp. 1-16). New York, NY: Guilford Press.

Jussim, L. & Eccles, J. S. (1992). Teacher expectations II: Construction and reflection on student achievement. *Journal of Personality and Social Psychology*, *63*(6), 947-961.

Kuklinski, M. R. & Weinstein, R. S. (2001). Classroom and developmental differences in a path model of teacher expectancy effects. *Child Development, 72*(5), 1554-1578.

Lanning, K. (1991). *Consistency, Scalability, and Personality Measurement*. New York, NY: Springer-Verlag.

Legislative Office of Education Oversight (1997). *Availability of minority teachers* (Report No. SP037-417). East Lansing, MI: National Center for Research on Teacher Learning. (ERIC Document Reproduction Service No. ED409902).

Macarov, D. (1978). Empathy: The charismatic chimera. *Journal of Education for Social Work, 14*(3), 86-91.

McAllister, G..& Irvine, J. J. (2002). The role of empathy in teaching culturally diverse students: A qualitative study of teachers' beliefs. *Journal of Teacher Education, 53*(5), 433.

Metz, M. H. (1990). How social class differences shape teachers' work. In M. W. Mclaughlin, J. E. Talbert, & N. Bascia (Eds.), *The context of teaching in secondary schools: Teachers' Realities* pp. 40-107). New York, NY: Teachers College Press

Noblit, G. W., Rogers, D. L., & McCadden, B. M. (1995). In the meantime: The possibilities of caring. *Phi Delta Kappan, 76*(9), 680. Retrieved from http://pdkintl.org/publications/kappan/

Noddings, N. (1992). *The challenge to care in schools*. New York: Teachers College Press.

Noddings, N. (1995). Teaching themes of care. *Phi Delta Kappan*, *76*(9), 675. Retrieved from http://pdkintl.org/publications/kappan/

Noddings, N. (2002). *Starting at home: caring and social policy*. Berkeley: University of California Press.

Patterson, C. H. & Purkey, W. W. (1993). The preparation of humanistic teachers for Schools of the next century. *Journal of Humanistic Education and Development,31*, 147-155.

Pierce, C. (1994). Importance of classroom climate for at-risk learners. *Journal of Educational Research, 88*(1), 37-42.

Reed, G. S. (1984). The antithetical meaning of the term 'empathy' in psychoanalytic discourse. In J. Lichtenberg, M. Bornstein, & D. Silver (Eds.), *Empathy I* (pp.7-24). Hillsdale, NJ: The Analytic Press.

Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Washington, D.C.: Economic Policy Institute.

Sandven, J. (1979). Social sensitivity as a factor in the teacher process. A theoretical Discussion and an experimental contribution. *Scandinavian Journal of Educational Research*, *23*(3), 131-150.

Stotland, E., Mathews, K. E. Jr., Sherman, S. E., Hansson, R. O., & Richardson, B. Z. (1978). *Empathy, Fantasy and Helping*. Beverly Hills, CA: Sage Publications.

Viadero, D. (2004, September 8). Declaration calls for more caring environments in schools. *Education Week*, 10.

Developmentally Responsive Teacher Practices Across the Middle-to-High-School Transition

Cheryl R. Ellerbrock
Bridget Abbas
Michael DiCicco
University of South Florida

Abstract

In this yearlong qualitative multi-site case study, researchers identified how eighth and ninth-grade teacher practices may support students' basic and developmental needs across the middle-to-high-school transition. Data were collected throughout 2009, including individual interviews, focus group interviews, observations, and artifact data of 23 participants. Findings suggest relational and academic teacher practices may help to meet students' needs across the transition but these practices were not consistent from one school site to the other and the responsiveness of these practices also varied across sites. Practices consistent with the warm demanding teaching stance may have promise for supporting students' needs during this period of schooling.

The middle-to-high-school-transition is classified as "the most difficult transition point in education" (Southern Regional Education Board [SREB], 2002, p. 24). Many students experience difficulty during this transition as they attempt to adjust to their new school environment (Cauley & Jovanovich, 2006; Cushman, 2006; Mizelle, 2005; Queen, 2002). The nature of school transitions, including the responsiveness of the sending and receiving school environments and teachers in these environments, play a pivotal role in supporting students' needs during this time (Eccles & Midgley, 1989; Eccles & Roeser, 2011; Eccles et al., 1993). Responsive teacher practices of eighth and ninth-grade teachers may help to support students during this tumultuous period of schooling.

The middle-to-high-school transition is understudied (Barber & Olsen, 2004; Hertzog & Morgan, 1998; Wilcock, 2007) and is typically examined at the middle or high school level only rather than in tandem or over time (e.g., Ellerbrock & Kiefer, 2010; Butts & Cruziero, 2005; Langenkamp, 2010). Over the last decade this transition has been receiving increased attention due to reform efforts aimed at improving secondary education (Felner, Favazza, Shim, Brand, Gu, & Noonan, 2001). Hertzog, Morgan, and Borland (2009) assert the middle-to-high-school transition is an ongoing process, not a single event, unfolding over time and across school sites. This perspective suggests transition research should be longitudinal and span across school sites, yet little research of this type is available. Further, few studies focus on educator and student perceptions of the middle-to-high-school transition and how educators can best support students' needs as they make the move to high school (see Akos & Galassi, 2004; Cushman & Rogers, 2008; Smith, Feldwisch, & Abell, 2006, for exceptions). The current study responds to Hertzog and colleagues' (2009) claim that the transition is a process unfolding over time. Further, this study addresses the aforementioned gaps in the literature by investigating eighth and ninth-grade

teacher practices that may support young adolescents' basic and developmental needs across the middle-to-high-school transition through the perspectives of educators and students.

The Need for a Developmentally Responsive Transition

Three conceptual frameworks undergird this study: self-determination theory (Deci & Ryan, 1985, 2000), Noddings' (2005) and others conceptualization of care in school, and stage-environment fit theory (Eccles & Midgley, 1989; Eccles & Roeser, 2011; Eccles et al., 1993). Self-determination theory advocates the need to support students' basic psychological needs for relatedness, competence, and autonomy in order to promote psychological growth, well-being, and motivation. Relatedness is defined as developing a sense of security and connection within a larger social network (Baumeister & Leary, 1995; Urdan & Schoenfelder, 2006). Feeling cared for by those within the school environment, including teachers, is critical to supporting student personal and academic growth (Noddings, 2005). Competence involves feeling able to successfully interact in one's social world (Deci & Ryan, 2000). Autonomy is the ability to self-regulate and the perception of being in control of one's actions and achievements (Deci & Ryan, 2000; Urdan & Schoenfelder, 2006). Satisfying these needs is crucial to understanding students' goal pursuits and attainment, including academic achievement (Deci & Ryan, 2000). Self-determination theory highlights the need for teachers to support students in their schooling by meeting these basic psychological needs.

Noddings' (2005) conceptualization of care in school suggests care is a fundamental need and feeling cared for by those within the school environment, including teachers, is necessary to support students' personal and academic growth. While difficult to define, care involves more than "a warm, fuzzy feeling that makes people kind and likable" (Noddings, 1995, p. 676) and "is a much more involved concept that cannot be defined solely by the presence of cordiality" (Schussler & Collins, 2006, p. 1465). It involves a sense of connection and support, is an important part of any relationship, and may not be accomplished without action (Chaskin & Rauner, 1995; Hayes, Ryan, & Zseller, 1994). Care can be both relational and academic. Relational care is comprised of supporting students' cognitive and social-emotional development, motivation, and learning (Eccles & Roeser, 2011; Roeser, Midgley, & Urdan, 1996; Wentzel, 1997) and enhancing students' sense of connection and belonging to their school (Noddings, 2005; Osterman, 2000; Roeser et al., 1996; Schussler & Collins, 2006). Teachers who foster relational care and connect with students may foster high-quality teacher-student relationships that meet student needs for relatedness (Davis, 2006; Roeser, Eccles, & Sameroff, 2000; Schmakel, 2008; Wentzel, 1997). Academically, care involves an ongoing quest for competence (Noddings, 1995) and involves "having high expectations and rigorous standards, pushing students further than they might believe they can go, and supporting them as they try to accomplish their goals" (Nieto, 2010, p. 264). The insistence on academic excellence from all students is identified in the literature as a caring, responsive teaching practice (Corbett, Wilson, & Williams, 2002, 2005; Nieto, 2010; Ross, Bondy, Gallingane, & Hambacher, 2008) and one that may promote a successful middle-to-high-school transition (Queen, 2002; SREB, 2002).

Recently Eccles and Roeser (2011) along with other scholars (see Ellerbrock 2012; Ellerbrock & Kiefer, 2010; Barber & Olsen, 2004) have applied stage-environment fit theory to the middle-to-high-school transition and students' first year in high school. Stage-environment fit theory suggests a mismatch between students' developmental needs and school and classroom

environments (Eccles & Midgley, 1989; Eccles et al., 1993). This mismatch may result in negative outcomes such as, declines in motivation and engagement and dropping out of school (Eccles & Roeser, 2011). Research suggests motivation and academic achievement increase when school personnel place a strong emphasis on supporting the needs of adolescent learners during school transitions (Butts & Cruziero, 2005; Eccles & Roeser, 2011; Gutman & Midgley, 2000; NASSP, 2006). Together stage-environment fit theory, Noddings' (2005) conceptualization of care in school, and self-determination theory provide a strong and interrelated theoretical base that guides the present investigation.

Responsive Teacher Practices

Teachers who meet students' needs, including the need to be cared for, are particularly important to adolescents during school transitions (Eccles & Midgley, 1989; Eccles & Roeser, 2011). Teacher practices such as holding high expectations, challenging students academically, and insisting students engage in classroom activities that foster a positive learning environment promote students' feelings of competence (Eccles & Roeser, 2011; Fredricks, Blumenfeld, & Paris, 2004; Whitlock, 2006; Wilson & Corbett, 2001). Student-centered practices such as making learning relevant to life outside of school, establishing flexible deadlines, and providing specific feedback on academic tasks may help foster a learning environment that meets student needs for competence and autonomy (Bishop & Pflaum, 2005; Eccles & Roeser, 2011; Garza, 2007; Schmakel, 2008, Wentzel, 1997). Providing opportunities for students to interact and connect positively with peers may support students' need for relatedness and belongingness in the classroom (Deci & Ryan, 2000; Juvonen, 2006). Additionally, believing in students' academic potential, getting to know students personally, listening to students, showing a genuine interest in students' academic and social lives, modeling caring behavior, teaching to understanding, being available and willing to help, and providing constructive feedback and academic support are teacher actions that address students' need for care (Garza, 2007; Weinstein, 1998; Wentzel, 1997). The warm demanding teaching stance combines teacher practices that focus on caring relationships and academic success to support students (Bondy & Ross, 2008; Ross et al., 2008; Ware, 2006). By focusing on the relational and academic needs of students, teachers can foster a climate of care in the classroom that supports student academic motivation and success (Wentzel, 1997).

Method

The aim of this yearlong qualitative multi-site case study was to gain a detailed understanding of how eighth and ninth-grade teacher practices may support young adolescents' basic and developmental needs across the middle-to-high-school transition. This study was part of a larger qualitative study that investigated the developmentally responsive nature of the transition from middle school (eighth grade) to high school (ninth grade). The following research question guided our investigation, "In what ways may eighth and ninth-grade teacher practices meet students' needs across the middle-to- high-school transition?" As a result, this case study features the perspectives of core subject teachers, students, and site-based principals regarding the ways teacher practices may support students' needs across the transition.

Context

Ford Middle School and Westshore High School are located in one of the 10 largest districts in the United States servicing a diverse ethnic and socio-economic student population. Ford's student enrollment during 2008-2009 was approximately 1559 students, including 480 eighth-grade students. Minority students made up 60% percent of Ford's population and 53% of students received free/reduced lunch. During the 2009-2010 school year, Westshore had 557 ninth-grade students and a total enrollment of approximately 1957 students. Fifty-four percent of Westshore's population was minority and 42% of students received free/reduced lunch. The demographics of these schools are representative of the school district (56.5% minority and 54% free/reduced lunch). Approximately 95% of Ford's 480 eighth-grade students matriculate to Westshore for high school.

Identification of Participants

A total of 23 people participated in this investigation: four students, four middle school teachers, 13 high school teachers, one middle school principal, and one high school principal (Table 1). Using purposeful sampling (Patton, 2002), the middle school principal selected one eighth-grade team to be part of this study based on a set of predetermined criteria (i.e., an interdisciplinary eighth-grade team, student population of the team represented the overall school demographics, and all team teachers were willing to participate). All 56 eighth-grade students on the team were invited to participate in order to capture "thick descriptions" (Ryle, 1949) from students and their eighth and ninth-grade teachers. A target sample of four students representative of the overall school demographics was selected. The student sample size is appropriate given the nature of the extensive, year-long longitudinal design, including shadowing each student throughout his or her day on multiple occasions. All four middle school team teachers who taught students' core subjects (math, science, language arts, and social studies) and 13 high school teachers who taught students' core subjects (math, science, English, social studies, reading, and freshman focus) participated. The larger number of high school teachers involved in the study is due to the lack of teaming at the high school level and the decision to include all students' core ninth-grade teachers. Years of teaching experience at the middle level varied from two years (Ms. Hamilton) to nine years (Mrs. Copeland) and one year (Mr. Oscar) to 33 years (Mrs. Peters) at the high school level. Mrs. Cramer, in her third year as the principal of Ford, and Mrs. Mauch, a 27-year veteran at Westshore, both participated. All adult participants and parents/guardians of student participants signed informed consent forms. Students were asked to provide written consent and verbal assent at the beginning of each semester.

Data Collection

This study utilized Stake's (2006) qualitative multi-site case study methodology. A multi-site case study approach is noted to be particularly appropriate when the intent is to gain a deeper understanding of participants' lived realities across multiple bounded systems, such as the two schools highlighted in the current investigation (Stake, 2006). Data

collection occurred throughout 2009 during the last semester of eighth-grade, summertime, and first semester of ninth-grade. Multiple data collection sources were utilized, including individual and focus group interviews, observations, and the collection of artifact data (Table 2).

A total of 23 individual and focus group interviews took place, including two teacher focus group interviews, two student focus group interviews, nine individual teacher interviews, eight individual student interviews, one middle school principal individual interview, and one high school principal individual interview. One student focus group interview was conducted in the spring semester of eighth grade and another in the fall semester of ninth grade. Individual student interviews took place toward the end of each semester after the focus group interview to extend on insights shared. Ninth-grade student individual interviews spanned across two days as a result of school-related events that shortened the duration of each class period. Students' core middle school and high school teachers were either individually interviewed or interviewed as a group, depending on their schedules. Middle school teachers (Mrs. Copeland, Ms. Mirabelle, Ms. Hamilton, and Ms. O'Connell) engaged in one focus group interview during their common planning period. Two dates were selected for high school teachers to participate in a focus group interview before school. Three teachers (Mr. Manns, Ms. Peters, and Mrs. Walters) participated in one high school teacher focus group interview. Due to a multitude of schedule conflicts that limited teachers' ability to participate in a focus group interview before school (e.g., tutoring, other meetings, child care issues), individual interviews were scheduled with nine teachers during the school day (Mrs. Matingly, Mr. Oscar, Mr. George, Mr. Leonard, Ms. Hines, Mrs. Cartright, Mr. Crespo, Mr. Roberts, and Mr. Simms). One high school teacher, Mrs. Erickson, was not interviewed due to multiple scheduling conflicts; however, she was included in observational and artifact data collection. Both the middle and high school principal participated in one individual interview. Semi-structured interview protocols were utilized. A student sample interview question was, "Describe how your teachers helped you with the transition into high school." A teacher and principal sample question was, "What are your expectations for academic excellence?" All interviews were audiorecorded and completely transcribed (141 single-spaced pages).

A total of 74 hours of observations took place (24 hours at the middle level and 50 hours at the high school level) where a researcher shadowed each student throughout his/her day at least twice at both levels. In eighth grade, students' team classes and other aspects of the school day (e.g., lunch, in-school activities) were observed. One event was observed over the summer (i.e., high school parent open house). In ninth grade, students' core classes along with other parts of the school day (e.g., team activities) were observed. Artifact evidence (e.g., students' schedules, syllabi, classroom rules, class assignments) were collected and used to help triangulate data. Multiple methods were utilized to ensure confidentiality and limit the effects of subjectivity, including the use of pseudonyms for participant and school names and member checks through which participants confirmed their perceptions were accurately represented.

Data Analysis

The present study utilized Hatch's (2002) inductive approach to data analysis. Most qualitative research is analyzed inductively for multiple reasons, including its ability to work flexibly within numerous qualitative paradigms and its ability to allow participant stories to surface by centering deeply on a particular entity (Hatch, 2002; Mayan, 2009). Like other inductive models, Hatch's version involves looking for patterns in data in an effort to generate general statements regarding the phenomena. After multiple reads of the complete dataset, researchers separated the data into analyzable parts, referred to by Hatch (2002) as frames of analysis, and compared frames to collaboratively form a consensus on which parts of the data require further analysis. These frames were then further analyzed to uncover domains that reflected semantic relationships (e.g., means-end semantic relationship such as X is a way to do Y; Figure 1). Emerging as the theme, relational and academic teacher practices may help to meet students' needs across the transition but these practices were not consistent from one school site to the other and the responsiveness of these practices also varied across sites. Teacher practices consistent with the warm demanding teaching stance may have promise for supporting students' needs during this period of schooling.

Results

Promoting Caring Teacher-student Relationships

Findings indicate eighth and ninth-grade teachers who engaged students in conversations personal in nature and who used team activities and in-class activities to connect with students helped to set the foundation for meeting students' needs and fostered teacher-student relationships grounded in care. While all middle school teachers used team activities and in-class activities to connect with students, only one high school teacher engaged in these practices. Fostering a sense of relatedness between teachers and adolescents is an essential element in nurturing a caring school environment and may help aid a responsive transition from one school to the next (Eccles & Roeser, 2011).

Conversations personal in nature. All eighth-grade teachers and 11 ninth-grade teachers in this investigation engaged in conversations personal in nature with their students. Such conversations were not directly related to academics and occurred in one-on-one, group, and whole class settings. Eighth-grade teachers attested to engaging in conversations personal in nature with individual students as well as with their entire class on a regular basis. They perceived such conversations allowed them to connect with their students and learn more about them. According to Mrs. Copeland, these conversations didn't have to be lengthy to be effective and could serve as a starting place for developing a deeper relationship with students:

> I think that just by the smallest contact too, they [eighth-grade students] want to build relationships with teachers. Even the smallest comment opens the door. You can be like, "Hey, I love your shoes" and they will be like, "Yeah, my mom bought them for me and then we went to the mall." And it's like a deluge and it's just the little comment that opens them up. They are like "Oh my gosh, she cares," and then they open up and talk.

Ms. Hamilton explained an instance when she purposefully made an attempt to connect with the entire class by asking about their time outside of school, "Guys, it's so great to see you back from spring break. So, for today's activity I want you to write about your spring break and then we are going to share it because they want to talk about it." She elaborated on the importance of such conversations, "It shows them that you want to know about their lives, what they do for fun, and their spare time." Ms. Mirabelle attested to the reciprocal nature of developing connections with students during such conversations, "They like talking. So, when you ask them a question, they open up and then they want to know something about you."

At the middle school level, eighth-grade teachers met as a team on a daily basis during a 90-minute common planning and lunch period to talk about students and their needs. These conversations often prompted personal exchanges with students. Ms. O'Connell explained, "Because we meet every day at lunch, we know that so-and-so was having a bad day this morning." She further described how she used the information acquired from her colleagues to express care toward her students, "They see you and you say, 'I heard that you were having a bad day this morning.' 'How did you know that?' …[S]o the students realize that we are working as a team to try and help them."

Students recognized their teachers had conversations about them. When reflecting back on her middle school experience, Lauren stated:

> I liked it better at Ford, because you felt more of a team. If you were having trouble in science, the teachers could talk because they all know you. The teachers could talk because they all know how you are doing in their classes.

During eighth grade, all four students reported their teachers knew them well and made efforts to engage them in whole class and one-on-one conversations personal in nature. For example, Katelyn recalled how one of her eighth-grade teachers engaged her in conversations revolving around her personal interests, "She [Ms. Hamilton] talks to me a lot…She knows that I like sports and I'm athletic." Lauren described how personal conversations with one of her eighth-grade teachers centered on a mutual experience, "Ms. O'Connell, the social studies teacher, is actually from Texas too. So a lot of the restaurants that we went to we can talk about and she knows where everything is." During classroom observations, all students were observed engaging in such conversations with all eighth-grade teachers on multiple occasions. For example, Jimmy talked with Ms. Copeland about her trip to the Grand Canyon. She shared with him her experience white water rafting on the Colorado River. He expressed excitement and asked for specific details about the event.

At the high school level, 11 teachers were observed engaging in conversations personal in nature with students and four of these teachers attested to participating in such conversations as a way to foster caring teacher-student relationships. On the first day of high school, Ms. Hines was observed asking students to fill out cards with information about their hobbies and lives outside of school. She explained how this information helped her initiate conversations with students, "I was able to say, 'Oh hey, I saw that you do this [on the information card]. That's pretty cool.'"

Additionally, Mr. George described how he utilized personal conversations with incoming ninth-grade students:

> In some of them I see fear, but I try to bring them out of their shells by talking to them. That's the only way. As soon as you do that, they relax. They relax. Then you can start talking to them and they begin to talk back.

Mr. Leonard stressed the importance of building relationships with students, "I make a connection and I do really well with rapport. If they understand your personality first, you build that trust by speaking, by not letting them fail, and expecting them to do work." In an attempt to help students foster positive relationships with their teachers, Mr. Matingly utilized class time to disclose his personal experiences as a secondary student:

> I thought that I was cool when I was in high school and that I would tell the teacher off. I would cuss teachers out and get suspended. I thought I was cool but I wasn't. It ruins your reputation.

During their ninth-grade year, all students referenced engaging in personal conversations with at least one of their ninth-grade teachers. Students used these conversations to inform their perceptions of teachers, including whether or not they like the teacher, could connect and relate with them, and approach them in time of need. As Jimmy stated, "Some teachers are really nice and some are really mean. The mean teachers just stay out and do their own thing and the nice teachers talk to the students and try to relate with the students." Similarly, on the second day of high school, Lauren claimed she already liked two of her teachers because they "seemed like they would be easy to talk to." Katelyn cited Ms. Walters as a teacher she connected with through talking about a mutual interest in sports, "She is interested in sports and reminds me of Ms. Hamilton [eighth-grade teacher] a bit." Jimmy reported connecting with his ROTC Sergeant, "…he's not like other teachers. If we do something bad, he doesn't make us stop. He talks to us like one of his friends."

Team activities and in-class activities. All eighth-grade teachers and one ninth-grade teacher in this investigation utilized team activities and in-class activities to connect and foster caring relationships with students. At the middle school level, two eighth-grade team activities, a time capsule and luau, fostered a sense of community and connection, demonstrating to students their eighth-grade teachers cared about them. For example, Lauren shared:

> I really liked the time capsule because we all got to put something in it, all the eighth graders on our team… Mrs. Copeland brought a rose bush, bought a gate to put around it, and brought a stone. I got to hold the stone and put it on the ground. So we all took a class picture of that. I think that by her taking the time to make all that stuff and get all that stuff for us was really special.

Near the end of students' eighth-grade year, the four eighth-grade teachers held a team luau to celebrate the conclusion of middle school. Students recognized the efforts teachers made to plan the luau (i.e., extra time, energy, and resources) remarking that such efforts were ways their teachers demonstrated care. Troy stated, "We had a party, a Hawaiian party, and the teachers

spent their own money to buy us hot dogs, watermelons, stuff like that." When asked how such efforts were perceived, he stated, "That they [teachers] care about us."

At the high school level, only one ninth-grade teacher was observed engaging in any sort of in-class activity with students purposefully designed to get to know students. This teacher, Mrs. Walters, spoke to students about the importance of knowing all classmates' names and working together as a cohesive group. She spent the first two days of the school year engaging in a class activity through which students got to know one another. Students were asked to come up with an adjective describing who they are that started with the same letter as their first name. At the start of the activity, she vowed to know all student names by the end of the second day of class, which she was able to accomplish. She told her students, "I will know your names by tomorrow, I promise." Classroom observations confirm that she knew all students' names by the second day of school. No other ninth-grade teachers were observed utilizing team activities and in-class activities nor did any other teachers describe using such experiences with their students. Similarly, no students in this investigation made reference to any additional ninth-grade teachers besides Mrs. Walters with whom they reported feeling a sense of connection.

Assuming Academic Responsibilities

During students' eighth-grade year at Ford and ninth-grade year at Westshore, teachers utilized numerous practices to assume academic responsibility as a way to prepare students for future responsibilities. Holding high academic expectations for students on both sides of the transition is noted as a transition best practice (SREB, 2002) and as an essential element in expressing care toward students (Bondy & Ross, 2008). Findings indicate middle and high school teachers assume academic responsibility differently. At the middle school level, teachers took ownership of academic responsibility for students, but high school teachers had the opposite view, insisting students take on all academic responsibility.

Teacher Responsibility. At the middle level, eighth-grade teachers insisted all students pay attention in class and complete all coursework to a quality level. Ms. Hamilton explained the eighth-grade teachers' attitude towards coursework, "It's not an option to put your head down and stare into space. It's not." She elaborated:

> For me, it goes with the high expectations thing. Zero is not an option in my class…I give them a 100 at the beginning of the semester and it's their job to keep it. I don't accept zeros. You have to do the work.

In addition to insisting students complete all coursework, three of the four eighth-grade teachers expected all students to complete coursework to quality. For example, on several occasions Katelyn and Jimmy were observed being told by Ms. Mirabelle to redo and resubmit their math homework because it was either not complete or not completed to quality. For these teachers, work not completed to quality was returned for revision and resubmission. Ms. Mirabelle explained, "With homework assignments that I collect and grade, if I don't think it's quality work, I don't grade it. I put a note on it that they have to redo it."

All eighth-grade teachers attested to and were observed being constantly "on" their students (i.e., constantly reminding students of what they need to do, insisting students pay attention in class, making directions overly explicit, providing multiple supports). Ms. Copeland stated, "You have to stay on them because they don't see the relevance [of the assignments]." Eighth-grade teachers expressed concerns that their practices were overly supportive, providing numerous academic supports that may not exist at the high school level. Ms. Hamilton stated, "I feel like we give them an abundance of safety nets and then they go next door [to Westshore High School] and they yank them." Although their intentions were to teach students' academic responsibility, eighth-grade teachers worried they may be overly involved and too supportive. Ms. O'Connell was concerned they may be hindering students' ability to fulfill the academic demands of high school, "I feel like it's [constantly being on them] babying them." It was her belief that high school students would not have "teachers that are really watching over them." She recounted telling her eighth-grade classes:

You are so lucky that someone is spelling it out for you, like 'Put your heading on your paper' because next year a teacher might get that [student paper] and put it in the garbage because there is no name on it.

One eighth-grade teacher, Ms. O'Connell, was observed vacillating between being constantly "on" her students like the other eighth-grade teachers and refusing to "baby" students like her high school counterparts. She explained, "I hold very high expectations in my classroom. If they are not prepared, I kick them out. …[T]o me it's important to teach them because in the workforce they are not going to get handouts."

Student Responsibility. At the high school level, ninth grade teachers emphasized the importance of meeting deadlines and actively resisted coddling students in an effort to prepare them for life after high school. All ninth grade teachers reported expecting students to complete coursework by the due date and would not accept late work. Their focus was on timely completion. Mrs. Cartright's policy regarding incomplete work was typical of the policies held by the ninth-grade teachers at the school, "Complete all assignments. Let me get serious here. If you have ten questions and you complete ten, you get 100%. If you have nine done and you give an excuse, you will get a zero." Mr. Oscar's policy was similar, "Late work, it is not going to fly. If you have late work and you turn it in, it is probably going to be a zero."

Students corroborated ninth-grade teachers' explanations of their strict classwork and homework policies. Troy described one such homework policy:

> Yeah, one of my teachers, as soon as the bell rings, if you don't have your book open and homework out, it's considered late. And if you don't turn in your homework on that day, the next day it's 50% off.

Similarly, Lauren stated:

> Some teachers are more strict or more lenient, but you have a certain time period where you have to get work done or you either get a zero or deducted points.

At the high school level, all ninth-grade teachers believed eighth-grade teachers shouldered too much student academic responsibilities. Ninth-grade teachers such as Ms. Hines spoke to the lack of "babying" in high school, "They [students] don't understand that they are now in high school and you are not going to get babied as much as you would in another place." Ms. Hines referenced the need to prepare students for a future in which they no longer have teachers to help them along:

> I think at the high school level there has to be a certain level of expectation and responsibility. I prod here and there. But if we are babying them throughout the whole high school then what skills are they going to have when they graduate?

Discussion

This study examined how eighth and ninth-grade teacher practices may help support students' needs across the middle-to-high-school transition. Three major conclusions were drawn. First, both eighth and ninth-grade teachers promoted caring teacher-student relationships as a way to meet students' needs across the transition. At both levels, teachers utilized informal conversations to promote a sense of relatedness and bolster connections with students. Eighth-grade teachers utilized team activities and in-class activities whereas only one such experience occurred at the high school level. Second, both eighth and ninth-grade teachers assumed academic responsibilities as a way to prepare students for future responsibilities; however, how eighth and ninth-grade teachers prepared students for future responsibilities were different. Third, teacher practices consistent with the warm demanding teaching stance may have promise for supporting students' needs during this period of schooling.

Although this study produced informative conclusions, it is not without limitations. While efforts were taken to aid in generalizability of the findings, due to this study's yearlong qualitative case study methodology including the voices of 23 participants, four of which were students, the ability to generalize the findings of this study is limited. More student participants might have yielded further insight into students' perceptions. Additionally, it is important to note data were collected during the second semester of students' eighth-grade year while data at the high school level were collected during the first semester of ninth-grade. Since data were collected at the end of the eighth-grade year, those teachers spent more time with students than ninth-grade teachers and, therefore, had increased opportunities to foster caring relationships and facilitate academic expectations that express care. Despite these limitations, this study provides a rich understanding of how eighth and ninth-grade teachers can support students' needs across the transition.

Literature on teacher care suggests caring teacher practices that promote a sense of connection and relatedness include: listening to students, showing a genuine interest in students and their academic and social lives, talking personally to students, and giving students a chance to talk (Garza, 2007; Hayes et al., 1994; Weinstein, 1998). These practices are echoed in the findings of this study. Middle school teachers fostered caring teacher-student relationships through the use of team activities and in-class activities. With the exception of Mrs. Walters's first day activity, such experiences were nonexistent at the high school level. Team activities at the middle school level, such as the luau and time capsule, fostered a sense of relatedness through deliberately promoting caring connections within the school setting that help to fulfill students' basic

psychological (Deci & Ryan, 1985, 2000) and developmental needs (Eccles & Midgley, 1989; Eccles & Roeser, 2011; Eccles et al., 1993). Findings indicate caring teacher practices, including conversations, events, and in-class activities, may support students' needs, especially their need for relatedness, across the transition (Ellerbrock & Kiefer, 2010; Bondy & Ross, 2008; Bosworth, 1995; Corbett, Wilson, & Williams, 2002, 2005; Garza, 2007; Nieto, 2010; Noddings, 2005).

Assuming academic responsibilities as a way to prepare students for future responsibilities may also be a way to support students across the transition. At the middle level, teachers assumed the responsibility for student learning, constantly pushing students to meet their level of expectation, even if it takes multiple attempts. With one exception, eighth-grade teachers exhibited the mindset advocated by Corbett et al. (2002), "If a task was worth giving, then it was worth doing: and if it was worth doing, then it was worth doing well" (p. 83). While eighth-grade teachers believed this mindset was necessary to academically support students, they were concerned that such a mindset may not adequately prepare students for high school. Eighth-grade teachers struggled with doing what they perceived was right for their students and doing what was needed to prepare students for high school teachers' academic expectations.

Academic expectations held by ninth-grade teachers centered on completion and timeliness. They adhered to a "real-world" mentality that centered on students' taking on complete academic responsibility and neither accepting late work nor offering second chances. Additionally, ninth-grade teachers did not discuss their expectations regarding the quality of completed assignments, just that all assignments needed to be complete on time. Additionally, they generally perceived eighth-grade teachers as taking on too much responsibility for students' academic success. Corbett et al. (2002) addresses the question of who is responsible for academics, students or teachers, and asserted until students value their education, adults are the responsible party. These discrepancies in the ways eighth and ninth-grade teachers facilitated academic expectations may not meet students' needs as they move from one school setting to another (Eccles & Midgley, 1989; Eccles & Roeser, 2011; Eccles et al., 1993).

Findings also suggest teacher practices consistent with those of warm demanding teachers may meet students' academic and relational needs. This teaching stance, in which teacher practices foster a responsive and caring teacher-student relationship and facilitate academic expectations, is important in promoting a responsive classroom environment (Bondy & Ross, 2008; Corbett et al., 2002, 2005; Ross et al., 2008; Ware, 2006), especially as students make the move from one school context to the next (Eccles & Midgley, 1989; Eccles & Roeser, 2011; Eccles et al., 1993). Findings indicate teachers on both sides of the transition have the potential to meet students' needs through the implementation of warm demander teacher practices; however, the warm demanding stance was not employed consistently across both levels. At the middle school level, three of the four teachers exhibited the warm demanding teaching stance by both promoting caring teacher-student relationships and assuming academic responsibilities. At the high school level, teachers exhibited warm demanding practices by promoting caring teacher-student relationships. Warm demanding academic teaching practices were not exhibited by any of the high school teachers.

This study highlights the importance of viewing the middle-to-high-school transition as a process spanning two time points, eighth and ninth grade, not a single event (Hertzog et al., 2009). While this yearlong longitudinal study underscores the need for examining how teacher practices on both sides of the transition support students' needs, attention must be paid to eighth and ninth-grade relational and academic practices in order to best meet students' needs across the transition. Additional research is needed to further investigate how teachers' use relational and academic practices to support students' needs and aid a responsive transition over a longer period of time. The exact role of care and how it may support students during this period of school merits increased attention. Future research on how the warm demanding teaching stance may help support a responsive transition is warranted. There is also a need to continue to listen to student voices in order to be responsive to their needs (Caskey, 2011) and uncover ways teachers may help support students' needs during the eighth and ninth-grade years as they make the transition from one school to the next.

References

Akos, P., & Galassi, J. (2004). Middle and high school transitions as viewed by students, parents, and teachers. *Professional School Counseling*, *7*(4), 212-221.

Barber, B. K., & Olsen, J. A. (2004). Assessing the transitions to middle and high school. *Journal of Adolescent Research, 19*(1), 3-30.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497-529.

Bishop, P. & Pflaum, S. (2005). Student perceptions of action, relevance, and pace. *Middle School Journal, 36*(4), 4-12.

Bondy, E., & Ross, D. (2008). The teacher as warm demander. *Educational Leadership, 66*(1), 54-58.

Bosworth, K. (1995). Caring for others and being cared for. *Phi Delta Kappan, 76*(9), 686-693.

Butts, M. J., & Cruziero, P. (2005). Student perceptions of factors leading to an effective transition from eighth to ninth grade. *American Secondary Education*, *34*(1), 70-80.

Caskey, M. (2011, November). *Important work ahead: Sustaining our vision and momentum.* William Alexander Memorial Lecture presented at the annual meeting of the American Middle Level Education (AMLE), Louisville, KY.

Cauley, K. M., & Jovanovich, D. (2006). Developing and effective transition program for students entering middle school or high school. *The Clearing House 80*(1), 15-25.

Chaskin, R. J., & Rauner, D. M. (1995). Youth and caring: An introduction. *Phi Delta Kappan, 76*(9), 667-674.

Corbett, D., Wilson, B., & Williams, B. (2002). *Effort and excellence in urban classrooms: Expecting—and getting—success with all students*. New York, NY: Teachers College Press.

Corbett, D., Wilson, B., & Williams, B. (2005). No choice but success. *Educational Leadership, 62*(6), 8-12.

Cushman, K. (2006). Help us make the 9th grade transition. *Educational Leadership,* 47-52.

Cushman, K., & Rogers, L. (2008). *Fires in the middle school bathroom: Advice for teachers from middle schoolers.* New York, NY: The New Press.

Davis, H. A. (2006). Exploring the contexts of relationship quality between middle school students and teachers. *The Elementary School Journal, 106*(3), 193-223.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York, NY: Plenum Press.

Deci, E. L., & Ryan. R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268.

Eccles, J. S., & Midgley, C. (1989). *Stage/environment fit: Developmentally appropriate classrooms for early adolescents.* In R.E. Ames & C. Ames (Eds.), Research on motivation in Education (Vol. 3, pp. 139-186). New York, NY: Academic Press.

Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence, 21*(1), 225-241.

Eccles, J. S., Wigfield, A., Midgley, C., Reuman, D., Mac Iver, D., & Feldlaufer, H. (1993). Negative effects of traditional middle schools on students' motivation. *The Elementary School Journal, 93*(5), 553-567.

Ellerbrock, C. R. (2012). Creating a family-like ninth-grade environment through interdisciplinary teaming. *Urban Education, 47*(1), 32-64.

Ellerbrock, C. R., & Kiefer, S. M. (2010). Creating a ninth-grade community of care. *Journal of Educational Research, 103*(6), 393-406.

Felner, R. D., Favazza, A., Shim, M., Brand, S., Gu, K., & Noonan, N. (2001). Whole school improvement and restructuring as prevention and promotion: Lessons from STEP and the Project on High Performance Learning Communities. *Journal of School Psychology, 39*, 177–202.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74,* 59–109.

Garza, R. (2007). "She teaches you like she were your friend": Latino high school students describe attributes of a caring teacher. *Journal of Border Educational Research*, *6*(1), 81-90.

Gutman, L. M., & Midgley, C. (2000). The role of protective factors in supporting the academic achievement of poor African American students during the middle school transition. *Journal of Youth and Adolescence, 29*(2), 223-248.

Hatch, J. A. (2002). *Doing qualitative research in educational settings.* Albany, NY: State University of New York Press.

Hayes, C. B., Ryan, A., & Zseller, E. B. (1994). The middle school child's perceptions of caring teachers. *American Journal of Education, 103,* 1-19.

Hertzog, C., & Morgan, P. (1998). Breaking the barriers between middle and high school: Developing a transition team for student success. *NASSP Bulletin*, *82*, 94-98.

Hertzog, J.C., Morgan, L., & Borland, K. (2009, November). *What research says to the practitioner about middle level to high school transition: Background and practices.* Concurrent session presented at the annual conference of the National Middle School Association (NMSA), Indianapolis, IN.

Juvonen, J. (2006). Sense of belonging, social bonds, and school functioning. In A. Winne & P. H. Winne, (Eds.), *Handbook of Educational Psychology* (pp. 655-674). Mahwah, NJ: Lawrence Erlbaum Associates.

Langenkamp, A. G. (2010). Academic vulnerability and resilience during the transition to high school: The role of social relationships and district context. *Sociology of Education*, *83*(1), 1-19.

Mayan, M. J. (2009). *Essentials of qualitative inquiry.* Walnut Creek, CA: Left Coast Press.

Mizelle, N. B. (2005). Moving out of middle school. *Educational Leadership*, 56-60.

National Association of Secondary School Principals. (2006). *Breaking ranks in the middle: Strategies for leading middle level reform*. Reston, VA: Author.

Nieto, S, (2010). *Language, culture, and teaching: Critical perspectives* (2nd ed.). New York, NY: Routledge.

Noddings, N. (1995). Teaching themes of care. *Phi Delta Kappan, 76*(9), 675-680.

Noddings, N. (2005). *The challenge to care in schools: An alternative approach to education: Advances in contemporary educational thought* (2nd ed.). New York, NY: Teachers College Press.

Osterman, K. F. (2000). Students' need for belonging in the school community. *Review of Educational Research*, *70*, 323-367.

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Queen, J. A. (2002). *Student transitions from middle to high school: Improving achievement and creating a safer environment.* Larchmont, NY: Eye on Education.

Urdan, T. & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology. 44,* 331-349.

Roeser, R., Eccles, J., & Sameroff, A. (2000). School as a context of early adolescents' academic and social-emotional development: A summary of research findings. *Elementary School Journal*, *100,* 443–471.

Roeser, R. W., Midgley, C., & Urdan, T. C. (1996). Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging. *Journal of Educational Psychology*, *88*(3), 408-422.

Ross, D., Bondy, E., Gallingane, C., & Hambacher, E. (2008). Promoting academic engagement through insistence: Being a warm demander. *Childhood Education, 84*(3), 142-146.

Ryle, G. (1949). *The concept of mind.* New York, NY: The University of Chicago Press

Schussler, D. L., & Collins, A. (2006). An empirical exploration of the who, what, and how of school care. *Teachers College Record, 108*, 1460-1495.

Schmakel, P. (2008). Early adolescents' perceptions on motivation and achievement in academics. *Urban Education, 43*(6), 723-749.

Smith, J. S., Feldwisch, R., & Abell, A. (2006). Similarities and differences in students' and parents' perceptions of the transition from middle school to high school. *Research in Middle Level Education*, *29*(10), 1-9.

Southern Regional Education Board. (2002). *Opening doors to the future: Preparing low achieving middle grades students to succeed in high school.* Atlanta, GA: Author.

Stake, R. (2006). *Multiple Case Study Analysis.* New York, NY: Guildford Press.

Urdan, T. & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology. 44,* 331-349.

Ware, F. (2006). Warm demander pedagogy: Culturally responsive teaching that supports a culture of achievement for African American students. *Urban Education, 41*(4), 427-456.

Weinstein, C. (1998). "I want to be nice but I have to be mean": Exploring prospective teachers' conceptions of caring and order. *Teaching & Teacher Education 14*(2), 153-163.

Wentzel, K. R. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology, 89*(3), 411-419.

Whitlock, J. L. (2006). Youth perceptions of life at school: Contextual correlates of school connectedness in adolescence. *Applied Developmental Science, 10*(1), 13-29. doi:10.1207/s1532480xads1001_2

Wilcock, A. (2007). Coping with high school--A transition for students and parents. *Primary & Middle Years Educator*, 26-31.

Wilson, B., & Corbett, D. (2001). *Listening to urban kids: School reform and the teachers they want.* Albany, NY: State University of New York Press.

Table 1.
List of Participants

Student Participants

Pseudonym	**Demographics**
Troy	Black male; eligible for free/reduced lunch
Jimmy	White male
Katelyn	White female; eligible for free/reduced lunch
Lauren	Hispanic female

Teacher and Administrator Participants

Pseudonym	**Position**	**Demographics**
Mrs. Copeland	Eighth grade science	White female
Ms. Hamilton	Eighth grade language arts	White female
Ms. Mirabelle	Eighth grade math	White female
Ms. O'Connell	Eighth grade social studies	White female
Mrs. Walters	Ninth grade English	White female
Mrs. Erickson	Ninth grade reading	White female
Mrs. Peters	Ninth grade reading	White female
Mr. Manns	Ninth grade math	White male
Mr. George	Ninth grade math	White male
Mr. Leonard	Ninth grade math	White male
Mr. Crespo	Ninth grade math	Hispanic male
Mr. Simms	Ninth grade math	White male
Ms. Hines	Ninth grade science	White female
Mr. Oscar	Ninth grade science	White male
Mrs. Cartright	Ninth grade social studies & freshman focus	White female
Mr. Roberts	Ninth grade social studies & freshman focus	White male
Mr. Matingly	Ninth grade freshman focus	White male
Mrs. Cramer	Middle school principal	White female
Mrs. Mauch	High school principal	White female

Table 2.
Data Collection Information

	Middle School Spring of eighth grade	**High School Fall of ninth grade**	**Total**
Focus Group Interviews	• 1 student focus group interview (*N*=4 students) • 1 core teacher focus group interview (*N*=4 teachers)	• 1 student focus group interview (*N*=4 students) • 1 core teacher focus group interview (*N*=3 teachers)	• 2 student focus group interviews (*N*=4 students) • 2 core teacher focus group interviews (*N*=7 teachers)
Individual Interviews	• 4 individual student interviews • 1 individual principal interview	• 4 individual student interviews • 9 individual core teacher interviews • 1 individual principal interview	• 8 individual student interviews • 9 individual core teacher interviews • 2 principal interviews
Observations	• 24 hours of student observations (*N*=4 students, each shadowed at least twice) • Team classrooms • In-school activities and events	• 50 hours of student observations (*N*=4 students, each shadowed at least twice) • Core classrooms • In-school activities and events	• 74 hours of student observations
Artifact Data	• Students' schedule • Syllabi • Classroom rules • Class assignments	• Students' schedule • Syllabi • Classroom rules • Class assignments	• Students' schedule • Syllabi • Classroom rules • Class assignments

Figure 1. Master Outline

Relational Domain
- Promoting teacher-student relationships at the eighth and ninth-grade level is a way to meet students' needs across the transition
  - Conversations personal in nature
  - Team activities and in-class activities

Academic Domain
- Assuming academic responsibility at the middle and high school levels is a way to meet students needs' across the transition
  - Teachers' responsibility
  - Students' responsibility

Figure 1. Master outline of relational and academic teacher practices that may help to meet students' basic and developmental needs across the transition.

Achievement Goals, Motivation to Learn, and Mathematics Anxiety among Pre-service Teachers

María D. Vásquez-Colina
Alyssa R. Gonzalez-DeHass
Joseph M. Furner
Florida Atlantic University

Correspondence related to this article should be directed to Dr. Maria D. Vasquez-Colina, Department of Educational Leadership and Research Methodology at Florida Atlantic University, (561) 297-2234 or mvasque3@fau.edu

Abstract

This paper reports findings of a pilot study examining the feasibility of a research design to investigate how achievement goals relate to the construct of math anxiety. In addition, we also consider how other important achievement-related behaviors, like self-efficacy, help-seeking, and self-regulation, might also relate to students' math anxiety. While math anxiety still remains a real issue affecting student performance and confidence, today it is more critical in our society with the greater emphasis on producing more students for careers in fields like Science, Technology, Engineering, and Mathematics (STEM). The total multiple regression model predicted a significant amount (43%) of the variation in math anxiety of participants.

Geist (2010) states that negative attitudes toward mathematics and what has come to be known as "math anxiety" are serious obstacles for young people in all levels of schooling today, and he feels that an anti-anxiety curriculum is critical in building students' confidence when working with mathematics especially in the light of a great push for more people going into the fields of Science, Technology, Engineering, and Mathematics (STEM). Helping students identify and address their math anxiety is critical in helping them cope with and overcome such anxiety that otherwise may negatively impact future choices in their academic and professional careers. As Boaler (2008) points out, it is critical to ensure students are confident and well prepared in mathematics if they are going to compete for such high-tech jobs today and in the future. Today, the United States is working to lead more young people into the STEM fields so we as a country can compete globally. Zollman (2012) believes that we need to evolve from learning for STEM literacy to using STEM literacy for learning to satisfy our societal, economic, and personal needs. If we are to build math confidence in our students, math teachers need to address head on the issue of math anxiety which often manifests itself as hesitancy or learned helplessness in observed math achievement.

Therefore, the purpose of this pilot study was to examine the feasibility of a research design intended to use in a larger-scale study examining how achievement goals relate to the construct of math anxiety and important achievement-related behaviors while learning math concepts.

Math Anxiety

Math anxiety may be defined as an *"...*inconceivable dread of mathematics that can interfere with manipulating numbers and solving mathematical problems within a variety of everyday life and academic situations" (Buckley & Ribordy, 1982, p. 1). The National Council of Teachers of Mathemtics (NCTM) (1989 & 1995b) recognized math anxiety as a problem and specifically included it in its assessment practices. Standard #10 (NCTM, 1989) prompts teachers to assess their students' mathematical dispositions; such as: confidence in using math to solve problems, communicate ideas, and reason.

As educators we need to know what causes this dread of mathematics so that it can be prevented and/or reduced. Causes of math anxiety may vary from socioeconomic status and parental background to the influence of teachers and the school system. Ahmed, Minnaert, Kuyper, &van der Werf (2012) examined the reciprocal relationships between self-concept and anxiety in mathematics. A sample of 495 grade seven students (51% girls) completed self-report measures assessing self-concept and anxiety three times in a school year. The analysis showed a reciprocal relationship between self-concept and anxiety in math (i.e., higher self-concept leads to lower anxiety, which in turn, leads to higher self-concept). Concluding that math self-concept and math anxiety are reciprocally related. Some educators believe that teachers and parents who are afraid of math can pass on math anxiety to the next generation, not genetically, but by modeling behaviors of their own discomfort with the subject. Research by Oberlin (1982) and Furner (1996) found that some teaching techniques actually cause math anxiety; (a) assigning the same work for everyone, (b) covering the book problem by problem, (c) giving written work every day, (d) insisting on only one correct way to complete a problem, and (e) assigning math problems as punishment for misbehavior.

Ineffective teaching practices are not the only cause of math anxiety. A student's lack of success with math may also be a cause of math anxiety and be heightened by any one of several factors: poor math instruction, an insufficient number of math courses in high school, unintelligible textbooks, or misinformation about what math is and what it is not. Many people often blame their failures on their lack of a mathematical mind, the notion that men are better than women at math, or that they have poor memories or learning disabilities. Sheila Tobias, an expert on the topic of math anxiety since the 1980's, contends that there are two myths about mathematics that need to be eliminated. One is that higher level math is too difficult for otherwise intelligent students to master, and another is that without mathematics you can live a productive intellectual and professional life (Tobias, 1993).

Willis (2010) gives over 50 strategies educators can use in any grade level to: (1) Rehabilitate negative attitudes about math; (2) Reduce mistake anxiety; and (3) Relate math to students' interests and goals. Having a better understanding of students' brains can help build foundational skills in math and other subjects and develop your students' long-term memory of academic concepts which can then prevent anxiety with mathematics.

A study by Perry (2004) indicated that 85% of students in an introductory college level math class claimed to have experienced anxiety when presented math problems. Jackson and Leffingwell (1999) showed another perspective in this study, with only seven percent of the

college students in their study not expressing math anxiousness. The prevalence of math anxiety in empirical studies is confounding; however, the effect of math anxiety is well documented. Even in populations of students where math is a foundational skill (e.g. engineering majors in college), researchers have found math anxiety to be present (Hembree, 1990; Ruffins, 2007). Sparks (2011) feels that as the STEM fields become more important for our students to study, our schools and teachers need to do more to address math anxiety so that our students are confident to study areas related to STEM. If math anxiety occurs frequently, then attention to the methods that are effective at overcoming math anxiety are important for teacher preparation as well as for in-service math teachers.

There are many things schools can do to help prevent math anxiety. Both teachers and parents play a critical role in helping to develop positive dispositions toward math. As with most intervention programs, early assessment and action help to develop positive math attitudes. The field of math education has recently made the push to increase and encourage math literacy, and along with that push has developed some useful materials to encourage math competence.

Reducing math anxiety is much different from preventing math anxiety. While every educator would like to prevent a student from experiencing math anxiety, some come to school already worried about being skilled at math. Ooten (2003) in her book, *Managing the Mean Math Blues*, outlines a four-step method for managing a persons' math anxiety. Ooten (2003) believes that a person who suffers from math anxiety needs to first lay the groundwork by coming to terms with their feelings and challenge their current beliefs and realize they are not alone; second, one must change their thoughts and negative thinking and use intervention strategies to improve one's thinking that they can be successful at math; third, one needs to know thyself, it is important that one knows his/her learning style/mode and that he/she applies approaches to doing math by successful people; and lastly fourth, once one has gained some confidence and strategies for doing mathematics they then must apply what they learned and actually do the math. All of Ooten's techniques require the teacher to first be aware and second to support the student in turning around their anxiety.

Math Anxiety among Preservice Teachers

Math anxiety exists in elementary education. Some studies have found that as high as 93% of preservice teachers express some anxiety, with anxiety beginning as early as the elementary school level (Jackson & Leffingwell,1999; Bekdemir, 2010; Dunkle, 2010; McAnallen, 2010; Blazer, 2011; Westenskow, & Moyer-Packenham, 2012). Many elementary education preservice teachers never took higher than a formal Algebra class or very many math classes as part of their studies to become a teacher (McAnallen, 2010). Some research has found that most elementary education majors reported having poor experiences with math courses in K-12 (Bekdemir, 2010; McAnallen, 2010). Other empirical studies discuss strategies that help in addressing math anxiety in preservice teachers like best practices such as cooperative groups, manipulatives, journal writing, and discussing feelings, clinical placements, etc. (Dunkle, 2010; Brown, Westenskow, & Moyer-Packenham, 2012).

Brown, Westenskow, and Moyer-Packenham, (2011) found that preservice elementary teachers' math anxiety may not necessarily always affect their math teaching anxiety. However, a

teacher's classroom behaviors, both overt and covert, during math instruction have proven to be influential on students' math anxiety (Jackson & Leffingwell, 1999). In addition, and critical to our line of research, is that a teacher's motivational goals towards learning can impact students' learning goals, self-efficacy, use of positive coping strategies, and willingness to seek help and ask questions (Butler & Shibaz, 2008; Friedel, Cortina, Turner, & Midgely, 2007). A topic we turn to in the coming sections.

Overcoming Math Anxiety

As can be seen from the above math anxiety literature, there are many motivational strategies that can be used to both reduce and prevent math anxiety which are in line with NCTM recommendations (1995a). Motivated learning strategies such as removing the importance of ego from classroom practice, emphasizing that everyone makes mistakes in mathematics and everyone has the capacity to improve, making math relevant, and letting students have some input into their own evaluations can be very useful in both preventing and reducing math anxiety and improving attitudes toward learning mathematics and then leading students to pursue a wider range of goals as it relates to future studies and career orientation. Synthesizing across this literature, it becomes apparent that many of these motivational strategies overlap with the research literature highlighting how influential a student's goal orientation can be to their learning. Therefore, of particular importance to this study was to look at research examining the relationships between preservice teachers' achievement goals and their levels of math anxiety. The next section introduces the goal constructs and reviews research that has examined how the goals that students' adopt relate to their anxiety and fear of failure with particular emphasis on those studies looking at math anxiety specifically.

Goal Orientation Theory

Goal orientation theory examines how the types of achievement goals students adopt are linked with important academic outcomes (Ames & Archer, 1988; Conroy & Elliot, 2004; Dweck & Leggett, 1988; Elliot, 2005; Elliot & McGregor, 2001; Ryan, Ryan, Arbuthnot, & Samuels, 2007; Sideris, 2008). Elliot and McGregor (2001) proposed a 2x2 conceptualization of achievement goals: performance approach goal, performance avoidance goal, mastery-approach goal and mastery-avoidance goal.

Students adopting performance goals are motivated to outperform others (performance-approach) or to avoid failure (performance-avoidance). Students adopting mastery-approach goals are positively motivated to master a task and advance one's learning; students adopting mastery-avoidance goals are negatively motivated to avoid misunderstanding and leaving a task un-mastered. While mastery-approach goals are related to deep processing, intrinsic motivation, and GPA; mastery-avoidance goals are related to disorganized studying, fear of failure, and test anxiety; performance-approach goals are linked to performance attainment but also sometimes to more surface-level processing; and performance-avoidance goals are positively linked to test anxiety and negatively linked to intrinsic motivation, exam performance, and GPA (Conroy & Elliot, 2004; Elliot, 2005; Finney, Pieper, & Barron, 2004; Furner & Gonzalez-DeHass; 2011; Ryan et al., 2007; Sideridis, 2008). Of particular interest to this paper is to tease apart how each of the goal constructs relates to math anxiety specifically.

Few published studies have actually examined the relationship between math anxiety and all four achievement goal constructs. Bong (2009) reported that elementary and middle- students with performance-approach, mastery-avoidance and performance-avoidance goals experienced more math anxiety (with the mastery-avoidance goal showing the strongest relationship with math anxiety). While mastery-approach goals appeared to be particularly beneficial, providing a stronger 'psychological armor', in combating adolescents' help-seeking avoidance and test anxiety in math classes. Putwain and Daniels (2010) examined how the relationships between test anxiety and competency beliefs are moderated by achievement goals among secondary mathematic students in England (ages between 11 and 12 years). Surprisingly, weak to moderate positive correlations were found between the test anxiety subscales of 'thoughts' (worrisome thoughts about a negative outcome for tasks) and 'autonomic reaction' (such as shaky hands during test-taking) and all four achievement goals, although effects were slightly stronger for both avoidance goals.

While not including all four goal dimensions in their analyses, a few additional studies have also examined the relationship between students' math anxiety and achievement goals using the earlier trichotomous framework (mastery goal and two performance goals). Middle school students' adoption of performance-avoidance or performance-approach goals was predictive of anxiety for math lessons (Skaalvik, 1997) and test anxiety in math (Middleton & Midgley, 1997). Zusho, Pintrich, and Cortina (2005) examined the relationships between the trichotomous framework of achievement goals and undergraduate students' motivational outcomes on math tasks and found that performance-avoidance goals were related to lower achievement scores, lower levels of competence, and higher levels of anxiety while mastery and performance-approach goals had positive outcomes for students' interest and competence perceptions.

Motivation to Learn

The achievement goal literature has uncovered that students' adoption of specific goals are linked with important motivational outcomes. Of particular interest to this study, are how achievement goals could be linked with students' self-efficacy, self-regulated learning, and help-seeking behavior and how this might shed light on students' math anxiety. In the trichotomous framework of achievement goals, mastery goals have been related to academic efficacy, self-regulated learning, deep processing, persistence and effort, help-seeking behavior, viewing learning tasks as a challenge and negatively predicting test anxiety; performance-approach goals have sometimes been linked to effort and persistence while studying, effective strategy use, and viewing learning tasks as a challenge, yet other times have been related to avoidance of help-seeking behaviors , and shallow or surface processing of information, and fear of failure and test anxiety (particularly in achievement situations that are perceived as a threat rather than an opportunity for challenge with little chance of failure); and performance-avoidance goals have been related to test anxiety, fear of failure, disorganized studying, surface processing, ineffective strategy use, procrastination, viewing learning tasks as a threat, and negatively related to task persistence, help-seeking, self-regulated learning, deep processing and academic efficacy (Cury, Elliot, Sarrazin, Da Fonseca, & Rufo, 2002; Elliot, 1999; Elliot & Church, 1997; Elliot & Harackiewicz, 1996; Elliot, McGregor & Gable, 1999; McGregor & Elliot, 2002; Middleton &

Midgley, 1997; Rawsthorne & Elliot, 1999; Shih, 2005). Thus, implying that achievement goal perceptions may also relate to student discomfort or anxiety where mathematics is concerned.

Additionally, when separating the mastery goal construct, Bong (2009) reported that mastery-approach goals appeared to be particularly beneficial, providing a stronger 'psychological armor', in combating adolescents' help-seeking avoidance in math classes (particularly those confronted with competitive, ability-focused learning environments). Students with mastery approach goals are interested in learning new skills and improving their understanding and competence; they are engaged in the process, not focused on the product. They are taking responsibility for their learning and engaged in activities that allow for self-regulation and self-direction. Their success is defined by individual improvement, they place value on effort, and their satisfaction is gained from working hard and learning something new.

Research Objectives

The first objective was to investigate the feasibility of the proposed research design including recruitment strategies using an internet-based instrument and the validity and reliability of the instrument. The second objective was to examine the nature of relationships between mastery and performance goals and math anxiety, self-efficacy, self-regulated learning, and help-seeking among pre-service teachers learning math concepts. Existing research investigating the relationships between students' achievement goals and reported math anxiety points to the benefits of students adopting mastery-approach goals and the deleterious effects of students adopting performance-avoidance goals. Slightly more inconclusive, and worthy of further study, are the effects of students adopting the performance-approach and mastery-avoidance goals.

Method and Sample

The pilot study sample consisted of 30 undergraduate students enrolled in sections of MAE 4350, Principles & Methods: K-9 School Math who were18 years old or older; 27 were females. All students taking the MAE 4350 course are either elementary education majors or elementary special education majors taking the course as a requirement for completing their program as elementary school teachers. The students had taken two other college level math classes as pre-requisites for this math methods course, one being at least College Algebra. The students took this class in their last two years of upper-division course work and had taken other education courses prior to this class and were familiar with how to write lesson plans. A non-random sample was used, and all students enrolled were given the chance to participate. While information about ethnicity was not collected, the university the students belong to is ranked as 32nd nationally in the number of bachelor's degrees conferred upon minorities (FAU, 2012).

Instruments

An online survey was distributed and stored electronically through *Google.doc*. The online survey was composed by four components: The Achievement Goal Questionnaire (AGQ) by Elliot and McGregor (2001), the Motivated Strategies for Learning Questionnaire (MSLQ) by Pintrich, Smith, Garcia and McKeachie (1991), the Abbreviated version of the Mathematics

Anxiety Rating Scale (MARS) by Alexander and Martray (1989) and additional demographic questions on age, gender and preferences of study.

AGQ. The AGQ was developed by Elliot and McGregor (2001) and included twelve Likert-type items. Students had to indicate if they thought the statement was very true (7) or not at all true of them (1). The AGQ scale was found to have strong reliability Cronbach's coefficients for four subscales: performance approach goals (.92), performance-avoidance goals (.83), mastery avoidance goals (.89) and mastery approach goals (.87). In the current study, Cronbach's alphas were as follows: performance approach goals (.74), performance-avoidance goals (.75), mastery avoidance goals (.87) and mastery approach goals (.82).

MSLQ. The MSLQ was developed by Pintrich, Smith, Garcia and McKeachie (1991). Information on predictive validity and internal reliability coefficients were available from this questionnaire. While the original full MSLQ has two main sections (a motivation section and a learning strategies section, totaling 81 items), only three subsections were chosen based on the purpose of study. Thus, the summarized MSLQ version in the current study included self-efficacy for learning and performance (from the motivation section) with eight items, metacognitive self-regulation (from learning strategies) with 12 items, and help seeking (from learning strategies) with four items. All questions were Likert-type items and the items ranged from 7(very true of me) to 1(not at all true of them). The MSLQ scale was found to have strong reliability Cronbach's coefficients for four subscales: self-efficacy learning and performance (.93), metacognitive self-regulation (.79), resource management: help seeking (.52) and mastery approach goals (.87). In the current study, Cronbach's alphas were as follows: self-efficacy learning and performance (.94), metacognitive self-regulation (.8) and resource management: help seeking (.5).

MARS. Originally developed by Alexander and Martray (1989), the Abbreviated Version of the Mathematics Anxiety Rating Scale (MARS) provides a measure of anxiety associated with math testing and numerical operations, and of math courses. It is an internally consistent and reliable 25-item scale (Alexander & Martray, 1989). Its primary purpose being to create an abbreviated version of the lengthier, 98-item MARS. The result was an internally consistent and reliable 25-item scale. The creators of the Abbreviated Version of the MARS used items from the actual full-scale MARS and from the Fennema-Sherman Mathematics Attitude Scales (Fennema, 1976) to develop the Abbreviated Version of the MARS. The abbreviated version is much like the original in that it is multidimensional. The Abbreviated Version of the MARS measures mathematics anxiety--specifically math test, number, and math course anxiety. The internal consistency of the scale scores using the 25 salient items calculating coefficient alpha was .96 for the 15 items measuring test anxiety, .86 for the 5 items measuring numerical task anxiety, and .84 for the 5 items that tested for math course anxiety. These coefficients compared favorably with the .97 coefficient alpha reported by Richardson and Suinn (1972) for the full-scale 98-item MARS. Two-week test-retest reliability of the abbreviated scale, based on the subsample of 62 students, was .86. The authors also stated, because math test anxiety has been identified consistently as the major component of math anxiety in other studies that the 25-item abbreviated MARS would seem to be more appropriate for students because of its efficiency, economy, and administrative ease in measuring math anxiety. Thus, the MARS scale used had 25 items

measuring the level of anxiety in an individual. All items ranged from 5(very much) to 1 not at all).

Procedures

Upon institutional review board (IRB) approval, recruitment information was sent to three course instructors to distribute in their classes. Students were invited to participate in an online survey made up of three existing instruments: AGQ, MLSQ sections and Abbreviated MARS. Instructors were asked to send two reminders via their Blackboard site and email.

The researchers asked if students would be willing to participate in the online survey. Students were assured that there was no penalty for not participating, that completion of the survey had no relation to their course grade, and that they could withdraw at any time. The survey did not show any student identifiers other than gender. Also, students were assured that data would only be reported in the aggregate. The students were directed by their instructor to a link that took them to the online survey.

Data Analysis and Results

Multiple regression analysis was used to predict math anxiety from a set of predictors. Due to the fact that math anxiety is likely to be influenced by many factors, such as achievement goals and learning strategies, two multiple regression models were conducted to examine the question if math anxiety (The Abbreviated Version MARS) could be predicted from achievement goals (AGQ) and motivated strategies for learning (MSLQ). In the current study, Cronbach's alphas for MARS, MSLQ, AGQ and their subscales ranged from .5 to .9, respectively. Overall, the means for the independent variables were high (Table 1). Conversely, the math anxiety mean indicated that students' level of anxiety were not that high.

Table 1
Descriptive statistics for dependent and independent variables

Variables	Mean	SD	*n*
Achievement goal	5.19	.83	27
Performance approach	5.12	1.48	27
Mastery avoidance	3.35	1.84	27
Mastery approach	6.69	.57	27
Performance avoidance	5.59	1.14	27
Learning strategies	4.91	.29	43
Math anxiety	2.64	1.02	27

SD = standard deviation

The first multiple regression model was used to discern the amount of variance that could be predicted from achievement goals and motivated strategies for learning scores in terms of math

anxiety. The total model predicted a significant amount (43%) of the variation in math anxiety which was significant, R^2=.4, $F(2,26)$= 9.15, p<0.01. Achievement goal contributed significantly to the model (β =.58, p <.01). A second regression model included only each of the four achievement goals as separate predictors. Specifically, this model discerned the amount of variance that could be predicted from performance approach, mastery avoidance, mastery approach and performance avoidance goal orientations regarding math anxiety. The results of the regression model indicated that four of the predictors explained 55% of the variance in math anxiety, R^2=.5, $F(4,26)$ =6.73, p<0.01). A positive correlation between math anxiety and two of the predictors, mastery avoidance and performance avoidance were found significant, r=.73, r=.45, p <.01. Mastery avoidance statistically predicted math anxiety model (β =.73, p <.01). Simple correlations were also calculated between the achievement goals and students' reported use of self-regulated learning strategies, self-efficacy, and help-seeking behaviors. Mastery approach and self-efficacy were slightly correlated, r = .36, p < .05. Mastery approach and self-regulation were moderately correlated, r=.65, p<.01. Performance avoidance and self-efficacy were negatively correlated, r=-34, p<.05.

Discussion

While these analyses are the result of a pilot sample, and any conclusions should be considered with caution, preliminary findings are discussed in reference to our proposed objectives. Study limitations include a small sample size and the fact that not all respondents answered all items.

Our first research objective was to examine the feasibility of the proposed research design including recruitment strategies using an internet-based instrument and the validity and reliability of the instruments. We found that the recruitment strategies and research design were appropriate for this type of study. As the use of technology increases to support conducting research, benefits and challenges in using such a tool appear as well. While participants have the option to not respond to questions, it seems that electronic surveys may increase the nonresponse rate. The main sources of error in electronic surveys include "sampling, coverage, nonresponse, and measurement error" (Couper, 2000, p.466). The pilot sample was selected from the target population which was preservice teachers in South Florida; therefore we used a representative group of participants to avoid sampling error considering that was a nonrandom sample. The fact that only a small number of respondents were not willing to complete all survey items created minimum nonresponse bias (Couper, 2000). The fact that the sample was representative, the data collection was successfully completed online and internal reliability was medium to high, proved that the research design was appropriate, thus a main study could be carried out with a larger sample.

Our second research objective was to examine the nature of relationships between students' achievement goals, math anxiety, self-efficacy, self-regulated learning, and help-seeking behaviors. We found that the model including achievement goals and motivated strategies for learning (self-efficacy, self-regulated learning and help-seeking) predicted a significant amount of the variation in math anxiety. Results seem to indicate that the achievement goals students adopt do relate to their level of math anxiety and other achievement behaviors and outcomes. The patterns of these relationships were expected in light of the literature as both avoidant type goals correlated with math anxiety, while a mastery approach orientation seemed to relate to

more positive achievement behaviors including students' self-regulated learning and self-efficacy. These findings suggest that continued investigation to examine these relationships with larger samples is worthy of further study which is forthcoming.

Summary and Final Thoughts

There is a real need for creating more young people who have a passion and interest in STEM fields, people who are confident in their abilities and will set goals to pursue careers in the area of mathematics and the sciences. The first step in such an important educational goal is to understand effective ways to reduce math anxiety and encourage more positive attitudes for learning mathematical concepts. If the goals students adopt have some relationship to beneficial achievement behaviors and a healthy outlook for learning math concepts, we can then consider how the research literature outlining suggestions for creating mastery-oriented classrooms may also help to reduce the anxiety students experience during mathematics instruction (Furner & Gonzalez-DeHass, 2011).

In particular, research in this area should continue to examine pre-service teachers' achievement goals and how they relate to their mathematics learning. The teacher's classroom behaviors, both overt and covert, during math instruction have proven to be influential on students' math anxiety (Jackson & Leffingwell, 1999). A teacher's achievement goals can impact students' learning goals, self-efficacy, use of positive coping strategies, and willingness to seek help and ask questions (Butler & Shibaz, 2008; Friedel, Cortina, Turner, & Midgely, 2007). What makes this study unique is that few published studies have actually examined the relationship between math anxiety and all four achievement goal constructs previously, and we did not locate any studies that examined pre-service teachers' goals and their relationship to math anxiety specifically.

Math anxiety is a very real thing affecting our future classroom teachers and their students. This study suggests that the achievement goals students adopt do relate to their level of math anxiety and other achievement behaviors and outcomes. The patterns of these relationships were expected in light of the research literature since both avoidant type goals correlated with math anxiety, while a mastery approach orientation seemed to relate to more positive achievement behaviors. It is critical in today's high tech STEM oriented society to encourage all young people, especially future classroom teachers, to take more mathematics and STEM related classes and perhaps come to terms with such anxiety before starting their career as a teacher. In this way, hopefully they will be able to better promote STEM-related fields with future students they may be working with. It is critical that we have both teachers and students on board having positive dispositions toward mathematics which then can promote more young people going into careers in the various STEM fields.

References

Ahmed, W., Minnaert, A., Kuyper, H., & van der Werf, G. (2012). Reciprocal relationships between math self-concept and math anxiety. *Learning and Individual Differences*, 22(3), 385-389.

Alexander, L., & Martray, C. (1989). The development of an abbreviated version of the mathematics anxiety rating scale. *Measurement and Evaluation in Counseling and Development, 22*, 143-150.

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology, 80*(3), 260-267.

Bekdemir, M. (2010). The pre-service teachers' mathematics anxiety related to depth of negative experiences in mathematics classroom while they were students. *Educational Studies in Mathematics*, 75(3), 311-328.

Boaler, J. (2008). *What's math got to do with it? Helping children learn to love their least favorite subject--and why it's important for America*. New York, NY: Penguin Group (USA) Inc.

Bong, M. (2009). Age-related differences in achievement goal differentiation. *Journal of Educational Psychology, 101*(4), 879-896.

Brown, A., Westenskow, A., & Moyer-Packenham, P. (2012). Teaching anxieties revealed: Pre-service elementary teachers' reflections on their mathematics teaching experiences. *Teaching Education, 23*(4), 365-385.

Brown, A. B., Westenskow, A., & Moyer-Packenham, P. S. (2011). Elementary pre-service teachers: Can they experience mathematics teaching anxiety without having mathematics anxiety?, *Issues in the Undergraduate Mathematics Preparation of School Teachers*, 5, 1-14.

Buckley, P. A., & Ribordy, S. C. (1982). *Mathematics anxiety and the effects of evaluative instructions on math performance*. Paper presented at the Mid-western Psychological Association, Minneapolis, MN.

Butler, R. & Shibaz, L. (2008). Achievement goals for teaching as predictors of students' perceptions of instructional practices and students' help seeking and cheating. *Learning and Instruction, 18,* 453-467.

Conroy, D. E., & Elliot, A. J. (2004). Fear of failure and achievement goals in sport: Addressing the issue of the chicken and the egg. *Anxiety, Stress, and Coping, 17*(3), 271-285.

Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly, 64*, 464-494.

Cury, F., Elliot, A., Sarrazin, P., Da Fonseca, D., & Rufo, M. (2002). The trichotomous achievement goal model and intrinsic motivation: A sequential meditational analysis. *Journal of Experimental Social Psychology, 38*, 473-481.

Dunkle, S. M. (2010). Remediation of math anxiety in preservice elementary school teachers (Doctoral Dissertation). Retrieved from ERIC. (ED517036)

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256-273.

Elliot, A.J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist, 34*(3), 169-189.

Elliot, A.J. (2005). A conceptual history of the achievement goal construct. In A. J. Elliot and C. S. Dweck's (Eds.). *Handbook of competence and motivation* (pp. 52-72). New York: Guilford Press.

Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 72,* 281–232.

Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology,70,* 461–475.

Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, *70*(3), 461-475.

Elliot, A. J., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*(3), 501-519.

Elliot, A.J. McGregor, H.A, & Gable, S. (1999). Achievement goals, study strategies, and exam performance: A mediational analysis. *Journal of Educational Psychology, 91*(3), 549-563.

Fennema, E. (1976). Fennema-Sherman mathematics attitudes scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. *Journal for Research in Mathematics Education*, *7*, 324-326.

Finney, S. J., Pieper, S. L., & Barron, K. E. (2004). Examining the psychometric properties of the achievement goal questionnaire in a general academic context. *Educational and Psychological Measurement, 64*(2), 365-382.

Florida Atlantic University-FAU. (2012). FAU's Diversity Initiative. Retrieved from http://www.fau.edu/hr/diversity/

Friedel, J. M., Cortina, K. S., Turner, J. C., & Midgley, C. (2007). Achievement goals, efficacy beliefs and coping strategies in mathematics: The roles of perceived parent and teacher goal emphases. *Contemporary Educational Psychology, 32,* 434-548.

Furner, J. M. (1996). *Mathematics teachers' beliefs about using the National Council of Teachers of Mathematics Standards and the relationship of these beliefs to students' anxiety toward mathematics.* (Unpublished doctoral dissertation). University of Alabama, Alabama

Furner, J.M. & Gonzalez-DeHass, A. (2011). How do students' mastery and performance goals relate to math anxiety? *Eurasia Journal of Mathematics, Science, & Technology Education, 7*(4), 167-182.

Geist, E. (2010). The anti-anxiety curriculum: Combating math anxiety in the classroom, *Journal of Instructional Psychology, 37*(1), p24-31.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21*, 33-46.

Jackson, C. D., & Leffingwell, R. J (1999). The role of instructor in creating math anxiety in students from kindergarten through college. *Mathematics Teacher, 92*(7), 583-586.

McAnallen, R. R. (2010). Examining mathematics anxiety in elementary classroom teachers (Published Doctoral Dissertation). University of Connecticut.

McGregor, H. A., & Elliot, A. J. (2002). Achievement goals as predictors of achievement-relevant processes prior to task engagement. *Journal of Educational Psychology, 94*(2), 381-395.

Middleton, M. J., & Midgley, C. (1997). Avoiding the demonstration of lack of ability: An underexplored aspect of goal theory. *Journal of Educational Psychology, 89*, 710–718.

National Council of Teachers of Mathematics. (1989*). Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (1995a). Mathematics anxiety [Supplemental Brochure]. Reston, VA: Author.

National Council of Teachers of Mathematics. (1995b*). Mathematics anxiety* [Supplemental Brochure]. Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics*. NCTM: Reston, VA.

Oberlin, L. (1982). How to teach children to hate mathematics. *School Science and Mathematics, 82*, 261.

Ooten, C. (2003). *Managing the mean math blues.* Upper Saddle River, New Jersey: Pearson Education.

Perry, A. B. (2004). Decreasing math anxiety in college students. *College Student Journal, 38*(2), 321-324.

Pintrich, P.R., Smith, D.A., Garcia, T., & McKeachie, W.J. (1991). A manual for the use of motivated strategies for learning questionnaire (MSLQ). National Center for Research to Improve Postsecondary Teaching and Learning. Ann Arbor: University of Michigan.

Putwain, D.W., & Daniels, R. A. (2010). Is the relationship between competence beliefs and test anxiety influenced by goal orientation? *Learning and Individual Differences, 20*, 8-13.

Rawsthorne, L. J., & Elliot, A. J. (1999). Achievement goals and intrinsic motivation: A meta-analytic review. *Personality and Social Psychology Review, 3*, 326–344.

Richardson, F. C., & Suinn, R. M. (1972). The mathematics anxiety rating scale: Psychometric data. *Journal of Counseling Psychology, 19*, 551-554.

Ruffins, P. (2007). A real fear. *Diverse Issues in Higher Education, 24*(2), 17-19.

Ryan, K.E., Ryan, A.M., Arbuthnot, K., & Samuels, M. (2007). Students' motivation for standardized math exams. *Educational Researchers, 36*(1), 5-13.

Shih, S. (2005). Role of achievement goals in children's learning in Taiwan. *Journal of Educational Research, 98*(5), 310–319.

Sideridis, G.D. (2008). The regulation of affect, anxiety, and stressful arousal from adopting mastery-avoidance goal orientations. *Stress and Health, 24*, 55-69.

Skaalvik, E. M. (1997). Self-enhancing and self-defeating ego orientation: Relations with task and avoidance orientation, achievement, self-perceptions, and anxiety. *Journal of Educational Psychology, 89*, 71–81.

Sparks, S. D. (2011). Math anxiety explored in studies, *Education Week, 30*(31), 1.

Tobias, S. (1993). *Overcoming math anxiety revised and expanded*. New York: Norton Publishing.

Willis, J. (2010). *Learning to love math: Teaching strategies that change student attitudes and get results*. Alexandria, VA: Association for Supervision and Curriculum Development.

Zollman, , A.(2012), Learning for STEM literacy: STEM literacy for learning, *School Science and Mathematics, 112* (1), 12-19.

Zusho, A., Pintrich, P.R., & Cortina, K.S. (2005). Motives, goals, and adaptive patterns of performance in Asian American and Anglo American students. *Learning and Individual Differences, 15*, 141-158.

Direct and Indirect Effects of Teacher Instruction and Feedback on Student Adaptive Help-Seeking in Upper-Elementary Literacy Classrooms

Carin Neitzel
University of Tennessee

Dennis Davis
Vanderbilt University

Abstract

Ninety-three fourth and fifth-grade students were observed once weekly for one semester during reading and writing instruction. A structured observational protocol was used to record information about instruction and feedback provided to these students by their teachers, as well as the students' participation, regulation, and self-instruction behaviors. A path analytic model was tested to investigate the direct and indirect effects of teacher instruction and feedback on student self-instructive information pursuits or adaptive help-seeking behaviors. The findings reveal that some aspects of teacher instruction and feedback influence student self-instructive information pursuits directly; however, the primary influence of teacher instruction occurs indirectly through effects on students' patterns of participation and monitoring. A student's personal pattern of participation and monitoring during teacher-directed instructional episodes was an important determinant of self-instructional information pursuits during independent literacy activities in the upper-elementary school classroom.

The transition from lower-elementary (K-2) to upper-elementary (3-5) is a time characterized by increased academic demands along with changes in the division of roles in the classroom. The renegotiation of responsibilities is especially apparent in reading and writing instruction. Upper-elementary students are expected to make an abrupt shift from instruction focused on decoding words to instruction focused on meaning-making and deep processing of texts (Chall, Jacobs, & Baldwin, 1990). After third-grade, the amount of time devoted to reading instruction significantly decreases, and students are expected to independently complete more lengthy reading and writing assignments (Durkin, 1974-1975). Not all students negotiate this transition successfully; consequently, even once successful students in lower-elementary may begin to struggle with the new academic expectations of the upper-elementary. It is important to understand the instructional practices and student learning behaviors and dispositions that may facilitate (or hamper) this transition and promote independent learning in reading and writing.

Self-regulated learners autonomously direct and manage their own learning instead of relying on continual guidance from external agents (Zimmerman, 1989). During teacher-directed instruction, self-regulated learners engage in the learning process through the employ of a variety of participation strategies. They answer questions, provide examples and make connections, generate solutions, and acquire and apply heuristics to help manage their cognitive loads. Self-regulated learners also are aware of what they know

and don't know (Winne, 1995; Zimmerman, 1995). When more information is needed to complete a task or master a concept, self-regulated learners use their teachers and peers as resources to acquire this information.

Upper-elementary teachers often expect students to have developed already a healthy repertoire of self-regulatory strategies. Although teachers may spend little time deliberately assisting their students along the path toward autonomous learning, even incidental interactions in the classroom may direct the course of students' development as effective independent learners. Research is needed that can assist educators in upper-elementary classrooms to better understand the development of self-regulation, the instructional practices influencing its development, and their critical role in the process. The purpose of this study is to investigate how information and feedback provided by teachers during literacy instruction relate to students' strategies for participation and regulation and eventual self-initiated and directed information pursuits.

Proposed Model of Adaptive Help-Seeking Information Pursuits

The goal of independent learning does not preclude use of external resources. A successful self-regulated learner initiates information pursuits about academic topics, tasks, strategies, or performances, among other participation options. These interactions are a forum for students to access, elaborate, organize, integrate, or verify information and enhance understanding (Butler & Winne, 1995; Woloshyn, Pressley, & Schneider, 1992). Self-instructional behaviors such as these are critical ingredients for literacy success. Therefore, understanding the factors that may contribute to the development of these self-instructional strategies is crucial. The present study seeks to test a model of students' self-instruction which hypothesizes that a student's use of an important component of self-instruction - information pursuit - is influenced by levels and types of instructional information and feedback available in the classroom as well student's own patterns of participation and regulation within instructional interactions with the teacher.

Help seeking, or information pursuit behavior, is an important strategy used by self-regulated learners. Whereas some types of help-seeking behaviors require excessive dependency on the teacher, *adaptive help-seeking* refers specifically to a set of behaviors that contribute to increased cognitive autonomy. *Adaptive help seeking* is a self-regulatory behavior used by a learner when he/she asks for assistance or information that is intended to facilitate his/her independent learning (Newman, 2002). Adaptive help seeking is a complex process involving specific competencies and resources, including: knowing what help is necessary, how to formulate an appropriate question, knowing who to ask for help and how to ask in a socially appropriate manner, willingness to admit that help is needed, and a classroom context that supports help seeking behavior (Newman, 2002). Help seeking behaviors in elementary school children vary depending on several social and contextual factors (Newman & Schwager, 1995).

Information made available by classroom teachers through direct instruction is a primary influential source in the development of academic self-regulation (Bronson, 2000; Neitzel & Stright, 2003; Scarr, 1992). Contingent on the content of these academic interactions,

students may begin to assemble a repertoire of general heuristics or learning strategies (Schunk & Cox, 1986), models or behavior scripts or routines (Schunk, 1987, 1997), and a set of performance standards and criteria (Zimmerman & Ringle, 1981); fundamental tools for self-regulation (Bugental & Goodnow, 1998; Winne, 1996). A student's beliefs about help seeking are influenced by the way teachers respond to help seeking behavior; in an effort to avoid negative evaluations, a student will act more passively when his/her help seeking is viewed negatively by a teacher (Newman, 2002). Adaptive help-seeking or informational pursuit is a strategy that can only be employed once a student has achieved a certain level of metacognitive competence; therefore, the questions and probes used and metacognitive information provided by teachers can help students become more skilled in recognizing gaps in their understanding and in asking questions (Newman 2002). Thus, teacher feedback plays two important roles in influencing help-seeking behavior. First, the type of feedback a student receives in response to his/her help-seeking attempts may determine the student's subsequent use of the strategy. Second, a student's exposure to teacher modeling of effective questioning may result in increased help-seeking competence.

Although certain invitations to participate may be extended during teacher-led instruction, students choose to engage in numerous ways. For example, they may simply recount rote facts and content information, forms of interaction that can be accomplished with only lower levels of information processing. Alternatively, students may choose to engage in the discussion of strategies, make personal connections and predictions, or self-evaluation or process assessments; forms of interaction that require deeper levels of information processing. In addition a student's active monitoring, either internally (e.g., independent, self-directed review of the accuracy of one's own work) or socially (e.g., comparing one's work to that of a peer), may be necessary to ensure appropriate and effective use of self-instructional skills. Before students can pursue additional information from the resources available in the classroom, they must perceive a gap between actual performance and desired performance (Butler & Winne, 1995; Nelson-Le Gall, 1985). We propose that students' select participation and regulation patterns ultimately may determine whether or not the students derive maximum benefit from social-contextual affordances of the classroom and function effectively with self-instructional tasks such as adaptive help-seeking pursuits.

Summary of Study

The ability to seek help appropriately is especially important for students in upper-elementary due to the increased cognitive demands placed on them during literacy instruction (Chall et al., 1990; Pressley et al., 1998). This study examines the relative contribution of a profile of social-contextual features of the classroom and students own participation and regulation strategies for understanding their subsequent self-initiated information pursuits in school. In sum, the central hypothesis of the study was that students' level of participation and information processing would mediate the relations between teachers' instructional information and feedback patterns and students' self-instruction strategies in elementary school literacy classrooms. Information gained from this study can enrich understanding of the development of these important tools for

execution and regulation of learning and offer much needed insight in the design of more supportive and effective literacy instructional practices.

Method

Participants

The data analyzed represent a subset of data collected in a larger study of self-regulation. An invitation to participate in this study was extended to all the students in four classrooms in two different schools in a southern region of the United States. One school served a predominantly African-American population in a low income urban neighborhood, and the other served a predominantly Caucasian population from a low income rural area. According to information from the U.S. Census Bureau quick facts and American Factfinder, just over 50% of the adults in these areas are high school graduates, less than 10% have bachelors degrees. The per capita income is between $14,000 and $16,000. Ninety-three students (25 fourth-graders and 68 fifth-graders) comprised the sample.

One of the classroom teachers taught two fifth-grade writing/reading classes in the urban school (a charter school in the neighborhood). Two of the remaining three teachers taught fourth-grade self-contained classes in the rural elementary school in a nearby town. The other teacher taught a fifth-grade self-contained class in the same rural school. Each of the teachers was fully certified in elementary education in the state of Tennessee; two of the four teachers had master's degrees in education. All four of the teachers were female.

Procedure

This study was conducted across the second semester of the school year. Teachers and their students were observed during reading/writing class instruction once a week throughout the semester. Each class was observed approximately 14 times and each observation lasted approximately one hour. Student academic behaviors and teacher instruction and feedback behaviors were coded for three-minute intervals using a structured observation protocol similar to that used by Neitzel in previous work (Neitzel & Stright, 2003). Each interval focused on four students located in the same area of the classroom. An observation rotation scheme was used to identify the target students for each three-minute observation interval. Each student was observed between four to six times during each observation period. This systematic rotation prevented the researcher from selecting students for an interval based on the quality of interaction he/she was having with the teacher. Teacher instructional behavior during each observation interval also was coded.

In addition, information about the instructional context was recorded for each interval. An interval was considered *teacher-directed* when the predominant direction of instruction was from teacher to student. Typically during these intervals, the teacher positioned herself at the front of the classroom, and students remained in their desks and raised their hands to ask questions or make comments. The context was coded as

independent when the students worked on an individual assignment, typically a worksheet or an assignment from a textbook. The *group* context code was used when the students were either assigned to groups by the teacher or formed their own groups. During group work, students typically worked on one task together as a group (i.e., only one product was expected), although there were instances of group work that involved individual work products.

Measures

Teacher Instruction and Feedback. The instructional information and forms of feedback available during teacher directed academic lessons were observed and assessed using frequency counts of each of the targeted instructional behaviors. Instructional information in the classroom was identified as either basic information (i.e., rote facts related to the content being studied) or metacognitive information (i.e., strategy and process information intended to enhance performance, including the identification of common pitfalls) and the frequency of each information type was recorded. An overall feedback score was computed to represent the total amount of performance feedback provided to each student. Additionally, frequency scores were calculated for two types of performance feedback: undeveloped (i.e., purely evaluative and corrective comments) and formative (i.e., evaluative comments that also included information about how to improve performance). Finally, information was obtained about the frequency of teacher questioning supportive of autonomous student thinking (i.e., questions that prompted students to generate strategies, reactions, and predictions).

Student Academic Behaviors. The students' academic participation, regulation, and information pursuit behaviors were assessed using an observational coding system based on an instrument developed in previous studies of students' academic self-regulation (Neitzel & Stright, 2003; Stright, Neitzel, Sears, & Hoke-Sinex, 2001). Adaptations were made to the coding system to facilitate assessment of students' specific participation and progress monitoring strategy use and information pursuits during teacher directed and student-centered instructional activities in the classroom. The students' participation, regulation, and information pursuit behaviors in each context were assessed using frequency counts of each of the targeted behaviors.

Three participation types were assessed: basic-level or concrete information contributions (i.e., recalling and recounting rote facts or content information), deep-level information contributions (i.e., providing strategies, making predictions or connections), and self or process evaluations. Two types of monitoring strategies also were tracked for each student: internal or self-monitoring (i.e., checking and correcting one's own work, independent review of one's own work) and external or social-monitoring (i.e., comparing one's work to that of peers). In addition to computation of overall participation and regulation frequency scores in each academic instructional context, composite scores were calculated for each student to represent the student's overall frequency of each of the levels of participation and types of monitoring in each of the instructional contexts.

The student's adaptive pursuits of information in the classroom were assessed using frequency counts of instances in which the student sought information in the classroom either from the teacher, a peer, or other resources in the classroom. The student's pursuits of three information types were coded: additional information, task-process information, and normative information. Instances of the student seeking objective, factual information, or requesting more details about a topic of study were counted as pursuits of additional information. Instances of the student seeking task or process (mastery) information, that included strategies, rationales, and explanations of procedures or how and why things work, were counted as task-process information pursuits. Instances of the student seeking information for the purpose of making social comparisons, evaluating performance in relation to peers, or determining relative standing ("Did I do it the way it's suppose to be?" "How did everyone else do it?" "Is mine as good as his?") were counted as normative information pursuits. Final scores for each type of information pursuit were calculated by averaging the frequency counts for each behavior from all observation periods in each of the instructional contexts.

Results

First, preliminary analyses were conducted in order to provide descriptive information about the teachers' instructional behaviors and the children's academic behaviors. In addition to the initial analyses reported here, the distributions of each variable were examined statistically and graphically and the assumptions associated with general linear model analyses were checked. All assumptions of normality, homogeneity of variance, and linearity were upheld. Next, a hypothesized model of the relations among the teachers' instructional behaviors and the students' academic participation and regulation behaviors during teacher directed lesson and information pursuits during child-centered learning activities in the classroom were explored through path analysis.

Teacher Instruction and Student Participation, Regulation, and Information Pursuits

Frequencies, ranges, means, and standard deviations were calculated for each of the teacher instruction and student participation and regulation variables (Table 1). During teacher directed instructional activities in the classroom, typically, the teachers asked directed questions (Mean = 3.02, SD = 2.80) and provided basic information (Mean = 2.25, SD = 2.04). The teachers also routinely asked questions that served to support the students' active thinking about the material (Mean = 1.20; SD = 1.15). However, the teachers less frequently provided information about task demands, strategies, or procedural management suggestions (Mean = .60; SD = 1.38). The teachers did not often provide performance information or evaluation (Mean = .44; SD = 1.30). During teacher directed instructional activities, the students participation was characterized most regularly by recounting information (Mean = .49; SD = .87) and self-monitoring (Mean = .40; SD = .77). Infrequently, the students' methods of participation involved more deep-level processing forms (Mean = .16; SD = .40) and social monitoring (Mean = .19; SD = .46). The students rarely engaged in evaluation (Mean = .04; SD = .17) during teacher-directed instructional activities in the classroom.

Frequencies, ranges, means, and standard deviations were calculated for each of the student academic information pursuit behaviors during student centered instructional activities in the classroom (see Table 1). The children most frequently pursued task process information during student centered academic activities (Mean = .41; SD = .24), but also regularly pursued normative or evaluative information (Mean = .22; SD = .14) as they worked without teacher direction.

Path Analyses to Test Model Predicting Students' Information and Help-Seeking Pursuits

The relations among teacher instruction and student participation and regulation behaviors during teacher-directed instructional activities and student information pursuits during student-centered academic activities in the classroom were investigated through path analysis, which is useful particularly when it is hypothesized that a variable or set of variables may be operating as mediators of the relations between variables (Hoyle & Smith, 1994). In such instances, path coefficients are more accurate measures of the relation between any two variables because the effects of the other variables in the model are controlled statistically (Duncan, 1975).

Standard model fitting procedures utilizing maximum likelihood estimation were performed using LISREL 8.2 (Joreskog & Sorbom, 1998). The overall fit of the hypothesized model was tested using the goodness of fit χ^2 statistic. A significant χ^2 indicates the model is significantly different than the data; consequently, a χ^2 with a probability greater than .05 indicates that the model adequately fits the data.

A full model, in which teacher instruction and feedback patterns have direct as well as indirect effects on students' self-instruction strategies (information pursuits), was hypothesized and investigated. The overall model fit the data well, χ^2 (df = 25, n = 69) = 23.37, $\underline{p}$ = .56. The adjusted goodness of fit index, which does not consider the number of estimated parameters relative to sample size, was .89 (desired value, > .90, according to Sorbom & Joreskog, 1982; Thompson, 2000) and lends support to the initial conclusion regarding overall model fit. The root mean square error of approximation (RMSEA), a measure of the discrepancy in model fit per degree of freedom, was .012 (target value, < .05; Byrne, 1998; Sorbom & Joreskog, 1982) indicating a good fit of the overall model.

In addition, the fit of each of the model components was evaluated by examining the path coefficients (reported in Figure 1), which represent magnitude of the unique relationship between variables (after relationships with the other variables are controlled). In the examination of model component contributions, several differences were noted in the magnitude of path coefficients compared to original bivariate correlations, an indication of the possible presence of interactions (Falk & Miller, 1992). Therefore, in these instances, potential interactions among teacher and student behaviors during teacher-directed activity in relation to students' information pursuit behaviors were tested using the following procedure. In a series of hierarchical regression equations predicting each of the information pursuit behaviors, main effects (teacher instruction and/or student participation/regulation factors) were entered first into the equation; and then in a second step, the interaction term was entered into the equation (Berry & Feldman, 1985; Lewis-

Beck, 1980). A significant change in R^2 between the two steps is an indication of significant non-additive effects (Pedhazur, 1982). In order to interpret the meaning of any confirmed interactions with the student information pursuit behaviors, low and high groups for each main effect (predictor) variable were created using median splits, relations with the relevant information pursuit behavior(s) were plotted, and follow-up tests of the interaction were performed using a procedure described by Aiken and West (1991) in which the strength of relationship for each group was tested to determine whether it was statistically significant. The results of the model component analyses and post-hoc follow-up analyses are reported below.

Teacher instructional information/feedback and student levels of participation/processing

Squared multiple correlations for this part of the model were examined to determine the amount of variance in student participation levels explained by the hypothesized relations with teacher instruction and feedback patterns. The profile of teacher instruction and feedback behaviors explained .35, .34, and .03 of the variance in students' recall, deep-level participation, and evaluation, respectively (evaluation was more highly related to students' own deep-level processing). Teacher instruction and feedback explained .23 and .27 of the variance in students' self-monitoring and social monitoring.

Specifically, performance information, formative and undeveloped, were related negatively to students' recall (β = -.29 and β = -.50, $p < .05$, respectively). In addition, foundational instruction (basic information) was related positively to student recall (β = .28, $p < .05$). Basic instructional information also was related positively to students' deep level participation (β = .28, $p < .05$). Performance information, undeveloped, was related negatively to student self-monitoring (β = -.26, $p < .05$). Basic instructional information and formative performance information were related positively to student social-monitoring (β =.37 and .31, $p < .05$, respectively), and autonomy supportive questioning was related negatively (β = -.22, $p < .05$).

Relations between teacher instruction/feedback patterns and student participation levels and students' self-instruction behaviors

The profile of teacher instruction and feedback behaviors along with student participation and processing levels explained .60 and .55 of the variance in students' normative information and procedural/task-process information pursuits. According to reduced model squared multiple correlations, teacher instruction and feedback uniquely explained .22 and .05 of the variance in students' normative and procedural/task-process information pursuits. In sum, it appears that teacher instruction and feedback is directly related to students' self-instruction behaviors; however, it most strongly contributes to students' informational pursuits indirectly, through its influence on students' patterns of participation levels and strategy use.

Normative Information Pursuits. Metacognitive information was the only aspect of teacher instruction to contribute uniquely (negatively) to students' normative information pursuits (β =-.25, $p < .05$). However, students' participation behaviors during teacher

directed activities were related significantly to their information pursuits during student led activities (independent work and group work with peers). Specifically, recall and evaluation were related positively and deep level participation was related negatively to students' pursuits of normative information (β = .62, .43, and -.48, p < .05, respectively). Social-monitoring also was associated with students' normative information pursuits (β =.35, p < .05). Follow up tests revealed that there also was a significant interaction between teacher instructional questioning and student participation forms as well as student social monitoring. In addition, follow up tests revealed an interaction between teacher metacognitive information and student participation. Students who engaged in high levels of recall were more likely to pursue normative information during student-directed work even if the teacher provided high levels of support for student autonomous thinking through their instructional questioning patterns (β = .43, p = .001); however, for students who engaged in lower levels of route recall, teacher questioning was related to decreased pursuits of normative information (β = -.23, p = .05). Although in general increased social monitoring was associated with increased normative information pursuits, in the presence of increased teacher instructional questioning supportive of student thinking, social monitoring was not associated with normative information pursuits (β = .02, p = .87). And, even students who engaged in higher levels of route recall were less likely to pursue normative information if the teacher provided higher levels of metacognitive information during teacher-directed instructional activities (β = -.34, p < .01).

Procedural or Task-Process Information Pursuits. Teachers' performance information (undeveloped) and autonomy supportive questioning were each uniquely related to students' pursuits of procedural or task-process information (β = .21, and .36, respectively, p < .05). In addition, students' self-monitoring was related positively (β = .29, p < .05) and recall was related negatively (β = -.38, p < .05) to procedural or task-process information pursuits. Follow up tests revealed that there also was a significant interaction between teacher instructional questioning and student monitoring. In addition, there was an interaction noted between teacher metacognitive information and student evaluation. Students who engaged in high levels of self-monitoring during teacher directed activity were not likely to pursue task-process information during student-directed work even if the teacher provided high levels of support for student autonomous thinking through their instructional questioning patterns (β = -.20, p = .08); in contrast, for students who engaged in lower levels of self-monitoring and higher levels of social monitoring, teacher questioning was related to increased pursuits of task-process information (β =.47, p < .001). Only students who engaged in higher levels of evaluation during teacher directed instruction appeared to benefit significantly from increased teacher metacognitive information (β = .51, p < .001).

Discussion

This study examined the relative contribution of a profile of instructional features in upper-elementary literacy classrooms and students own participation and regulation strategies for understanding their subsequent self-initiated information pursuits during

independent or student led literacy activities in school. The central hypothesis of the study was that students' level of participation and information processing would mediate the relations between teachers' instructional information and feedback patterns and students' self-instruction strategies in elementary school literacy classrooms. The findings suggest that some aspects of teacher instruction and feedback influence student self-instruction directly; however, the primary influence of instruction occurs indirectly through the effects of various instructional components on students' own patterns of participation and monitoring during teacher-led literacy instruction activities.

Self-regulated literacy learning is not exclusively a cognitive process – social components of literacy are at play as well (Freebody, Luke, & Gilbert, 1991; Gee, 1996; Street, 2003. In this study, the social component of interest was the student's ability and willingness to seek out information from social sources (i.e. peers and teacher) during reading activities. The findings of this study indicate that the social dynamics of the classroom (how students participation and monitoring are sanctioned) impact how students pursue information related to reading and writing activities that have been assigned. A student's personal pattern of participation and monitoring during teacher-led episodes was an important determinant of self-instruction during independent work and the relationship between instructional input and performance feedback provided during teacher-directed literacy instruction and students' self-instruction during subsequent episodes of independent work was best described by taking into account the students' personal repertoires of instructional management behaviors during teacher-led activities. In sum, even if a teacher provided high-quality instruction and extensive feedback, students did not necessarily benefit if the classroom culture had not provided ample opportunity for monitoring and deep level participation.

Numerous studies and reports have documented the role of strategies in reading and writing development (Graham & Perin, 2007; Panel, 2000; Pressley, 2002). Most of this research focuses on how students use strategies to interact with and comprehend texts. Researchers interested in cognitive strategies instruction emphasize that good readers are able to assess whether or not what they read makes sense (monitoring) and employ strategies to ensure comprehension (such as summarizing, drawing inferences, etc.). (Dole, Duffy, Roehler, & Pearson, 1991). This study expands the definition of strategic reading and writing to include strategies that are enacted socially (i.e., students' strategic information pursuits while working on reading and writing activities). We view reading and writing classroom as communities of practice (Lave & Wenger, 1991) in which the literacy practices that occur in these spaces are ideologically governed (Barton, Hamilton, & Ivanic, 2000).

Before they seek out information, students must first be aware of a gap in their knowledge. Such metacognitive awareness of their reading and writing abilities and literacy tasks is a necessary prerequisite for adaptive help-seeking or informational pursuits (Baker, 2002; Baker & Brown, 1984; Pressley, 2003). Metacognitive awareness allows students to evaluate their progress during a reading and writing activity and identify missing information that is needed to ensure success. However, although information-seeking behavior is an important literacy practice, it is one that is not always

supported in traditional reading and writing classrooms. Teacher instructional practices may promote or inhibit student adaptive help-seeking by providing access (or not) to important strategy, task, and performance enhancement information. In the absence of availability of this information from the teacher, students may be ill-equipped for the demands of the literacy classroom in upper-elementary school in which the amount of time devoted to reading instruction significantly decreases, and students are expected to independently complete more lengthy reading and writing assignments (Durkin, 1974-1975).

Before students seek out information from teachers and peers, they also must feel empowered to do so – the social dynamics of the classroom have to allow for this type of strategy use. Reading and writing activities have to be viewed as social practices instead of isolated cognitive accomplishments (Freebody, Luke, & Gilbert, 1991; Gee, 1996; Street, 2003). Purcell-Gates et al. (2004) recently stressed this point saying "We reject this implication that the social and the cognitive are independent and incommensurable…. We suggest that a more accurate way of envisioning the relationship between the sociocultural and the cognitive is as relating transactionally in a nested relationship, with the cognitive occurring within the sociocultural context." Intentionally or unintentionally, components of teacher instructional practices may sanction certain forms of participation and regulation. The classrooms in this study were mostly traditional in that students interacted more with worksheets and workbooks than with authentic texts (books, stories, nonfiction sources). In these classrooms, cognitive aspects of reading and writing development predominated – but even still, the social aspects were not totally absent. There were many instances observed in which students initiated social maneuvers for acquiring the information they felt they needed. Students found ways to seek out information socially; however, it was the teacher-sanctioned participation and monitoring during teacher-directed instruction that ultimately determined the extent to which information seeking occurred during student-centered literacy learning activities in the classroom.

The findings derived from this study holds important implications for both practitioners and researchers. Armed with a deep understanding of the ways in which their instructional feedback patterns help students learn to orchestrate their own learning, teachers can begin to facilitate self-regulatory development in their students. For researchers, the model may inform the way they think about the role of classroom experience and individual resources in shaping patterns of self-regulatory behavior. Additionally, this research provides evidence that a teacher's influence on his or her students goes beyond traditional measures of academic growth. An important component of educational quality, as defined in the context of this study, is the development of information pursuit habits that will help students independently engage in self-instruction in the literacy classroom.

In order to construct knowledge, learners must be actively engaged with information and activities available in classroom settings (McCaslin & Good, 1996). Participation is a critical tool for academic success (Schneider & Bjorklund, 1992). Reciprocally, information has a strong influence on cognitive engagement and forms of engagement

with tasks (Winne, 1995). In their early academic interactions, students may begin to assemble a repertoire of general heuristics or learning strategies (Schunk & Cox, 1986), models or behavior scripts or routines (Schunk, 1987, Schunk & Zimmerman, 1997), and performance standards and criteria (Zimmerman & Ringle, 1981); fundamental tools for self-regulation (Bugental & Goodnow, 1998; Winne, 1996), particularly subsequent self-instructional behaviors such as adaptive help-seeking or informational pursuits. Thus, students' subsequent information pursuits and the information options emphasized in their pursuits may be contingent on the content or forms of information made available as well as their early participation and management. However, the results of this study also illuminate the need to investigate perceived (and actual) instructional context purposes or constraints; as well as task relevant variation in "need," selection criteria and factors influencing choice, and the precursive or adaptive function of particular strategies

References

Baker, L. (2002). Metacognition in comprehension instruction. In C. C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices*. New York: The Guilford Press.

Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson, R. Barr, M. L. Kamil & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 353-394). New York: Longman.

Barton, D., Hamilton, M., & Ivanic, R. (2000). *Situated literacies: Reading and writing in context*. London: Routledge.

Bronson, M. B. (2000). *Self-regulation in early childhood: Nature and nurture*. New York: The Guilford Press.

Bugental, D. B., & Goodnow, J. J. (1998). Socialization processes. In E. M. Hetherington (Ed.), *Handbook of child psychology: Socialization, personality, and social development* (Vol. 4, pp. 389-462). New York: Wiley.

Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology, 78*(3), 210-216.

Butler, D. L. & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research,* 65(3), 245-281.

Chall, J. S., Jacobs, V., & Baldwin, L. (1990). *The reading crisis*. Cambridge, MA: Harvard University Press.

Cowie, B., & Bell, B. (1999). A model of formative assessment in science education. *Assessment in Education, 6*(1), 101-116.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of educational research, 58*(4), 438-481.

Dole, J. A., Duffy, G. G., Roehler, L. R., & Pearson, P. D. (1991). Moving from the old to the new: Research on reading comprehension instruction. *Review of Educational Research, 61*(2), 239-264.

Durkin, D. (1974-1975). A six year study of children who learned to read in school at the age of four. *Reading Research Quarterly, 10*(1), 9-61.

Freebody, P., Luke, A., & Gilbert, P. (1991). Reading positions and practices in the classroom. *Curriculum Inquiry, 21*(4), 435-457.

Gee, J. P. (1996). *Social linguistics and literacies: Ideology in discourse 2nd edition*. Philadelphia: Falmer.

Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools*. New York: Carnegie Corporation of New York.

Heath, S. B. (1988). Protean shapes in literacy events. In E. R. Kintgen, B. M. Kroll & M. Rose (Eds.), *Perspectives on Literacy*. Carbondale, IL: Southern Illinois University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York, NY: Cambridge University Press.

National Reading Panel. (2000). *Teaching children to read*. Washington, DC: National Institute of Child Health and Development.

Natriello, G. (1987). The impact of evaluation processes on students. *Educational psychologist, 22*(2), 155-175.

Neitzel, C., & Stright, A. D. (2003). Relations between mothers' scaffolding and children's academic self-regulation: Establishing a foundation of self-regulatory competence. *Journal of Family Psychology, 17,* 147-159.

Newman, R. S. (2002). How self-regulated learners cope with academic difficulty: The role of adaptive help seeking. *Theory into practice, 41*(2), 132-138.

Newman, R. S., & Schwager, M. T. (1995). Students' help seeking during problem solving: Effects of grade, goal, and prior achievement. *American educational research journal, 32*(2), 352-376.

Perry, N. E., VandeKamp, K. O., Mercer, L. K., & Nordby, C. J. (2002). Investigating teacher-student interactions that foster self-regulated learning. *Educational psychologist, 37*(1), 5-15.

Pressley, M. (2002). Comprehension strategies instruction: A turn-of-the-century status report. In C. Collins & M. Pressley (Eds.), *Comprehension Instruction: Research-Based Best Practices*. New York, NY: Guilford Press.

Pressley, M. (2003). Metacognition and self-regulated comprehension. In A. Farstrup & J. Samuels (Eds.), *What research has to say about reading instruction* (pp. 291-309). Newark, DE: International Reading Association.

Purcell-Gates, V., Jacobson, E., & Degener, S. (2004). *Print literacy development: Uniting cognitive and social practice theories*. Cambridge, MA: Harvard University Press.

Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Development, 63,* 1-19.

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research, 57,* 149-174.

Schunk, D. H., & Cox, P. D. (1986). Strategy training and attributional feedback with learning disabled students. *Journal of Educational Psychology,78,* 201-209.

Schunk, D. H. & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist,* 32(4), 195-208.

Street, B. (2003). What's "new" in New Literacy Studies? Critical approaches to literacy in theory and practice. *Current Issues in Comparative Education, 5*(2), 77-91.

Winne, P. H. (1995). Inherent details in self-regulated learning. *Educational Psychologist, 30*(4), 173-187.

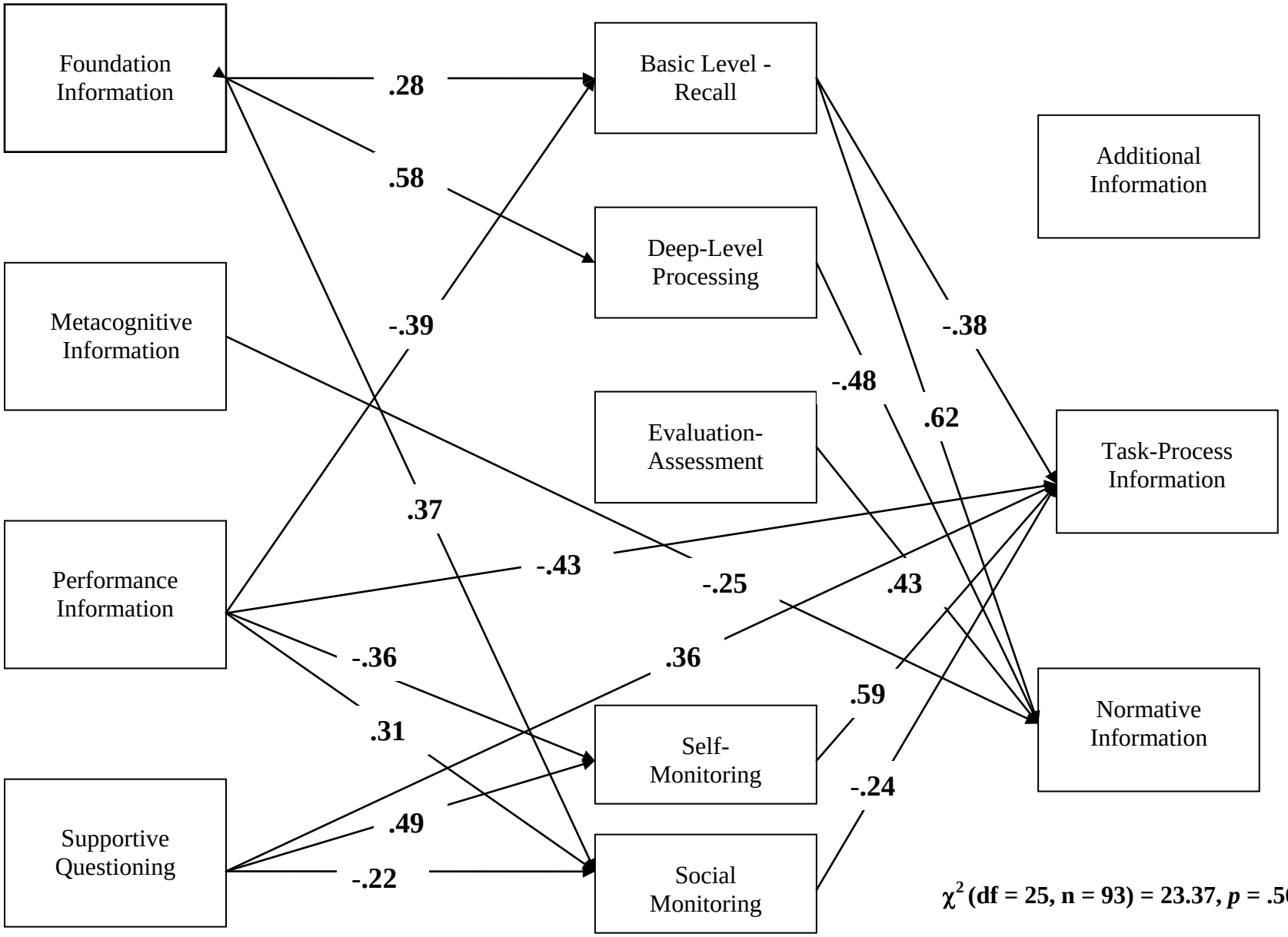
Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences, 8* (4), 327-353.

Woloshyn, V. E., Pressley, M. & Schneider, W. (1992). Elaborative-interrogation and prior-knowledge effects on learning of facts. *Journal of Educational Psychology*, 84(1), 115-124.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81*(3), 329-339.

Zimmerman, B. J. (1995). Self-regulation involves more than just metacognition: A social cognitive perspective. *Educational Psychologist, 30*(4), 217-221.

Zimmerman, B. J., & Ringle, (1981). Effects of model persistence and statements of confidence on children's self-efficacy and problem solving. *Journal of Educational Psychology, 73,* 485-493.


Foundation Information
Metacognitive Information
Performance Information
Supportive Questioning
Basic Level - Recall
Deep-Level Processing
Evaluation-Assessment
Self-Monitoring
Social Monitoring
Additional Information
Task-Process Information
Normative Information
.28
.58
-.39
.37
-.43
-.36
.31
.49
-.22
-.25
.36
-.48
.62
-.38
.43
.59
-.24
χ^2 (df = 25, n = 93) = 23.37, *p* = .56

Elementary General and Special Education Teachers' Mathematics Skills and Efficacy

Margaret M. Flores
Auburn University

Jennifer Thornton
University of Texas at San Antonio

Toni M. Franklin
Columbus State University
Vanessa M. Hinton
Auburn University

Shaunita Strozier
Valdosta State University

Abstract

The purpose of this study was to extend the literature regarding elementary teachers' beliefs about mathematics instruction to include special education teachers by surveying special education and general education teachers' mathematics teaching efficacy. In addition, the researchers' surveyed teachers' mathematics skills. The participants (n=178) were pre-service elementary level general and special education teachers enrolled in two major state universities in the Southern United States. The participants completed surveys of K-6 mathematics content and completed the Mathematics Teaching Efficacy and Beliefs Instrument (MTEBI). A MANOVA was conducted to investigate the differences between pre-service general and special education teachers' mathematics computation skills, problem solving skills, personal mathematics teaching efficacy, and mathematics teaching outcome expectancies. Findings indicate differences in participants' outcome expectancies and problem solving performance. However, both groups of pre-service teachers performed similarly in the areas of computation and teaching efficacy. The results have implications for teacher preparation.

The ability to demonstrate mathematical skill is critical for individuals' success, accounting for variances in employment, income, and work productivity more so than intelligence and reading ability (Rivera-Batiz, 1992). Early mathematics skills lay the foundation for advanced mathematics performance (Houchins, Shippen, & Flores, 2010). Therefore, it is essential that all children receive quality mathematics instruction in the early grades. In an effort to promote quality instruction for all students, reform efforts such as implementation of the National Council of Teachers of Mathematics Standards (NCTM, 2000), the *No Child Left Behind Act* (2002), and the *Individuals with Disabilities Education Improvement Act* (2004) have emphasized standards for practice, promoted evidence-based instruction, and progress monitoring. Additionally, the National Mathematics Advisory Panel (2008) calls for continued improvement of pre-service teachers' knowledge and pedagogical skills in mathematics.

A framework that has been adopted by many states as a way of meeting the challenge of reform efforts is response to intervention (RTI). RTI is a multi-tiered intervention model that involves

an interdisciplinary approach involving both general and special education. This model involves implementation of evidence-based practices and ongoing progress monitoring (Fuchs, 2003; Harris-Murri, King, & Rostenburg, 2006). RTI consists of instructional support for learning and is provided in levels called tiers which help students achieve educational goals, and become more specific and intensive depending on students' needs. RTI typically consists of three tiers of instruction (Hoover & Patton, 2008). Tier one is evidence-based core instruction that all students receive with progress monitoring occurring approximately three times a year. In tier one, all students, are provided evidence-based instruction in a general education classroom. Students who do not make adequate progress within the context or tier one receive tier two instruction. Tier two instruction is targeted interventions that address students' needs, provided through differentiated instruction within small groups of students in the inclusive setting by a general education teacher. Students who do not make progress in tier two are provided intervention support known as tier three. Tier three interventions are intensive, individualized, and are provided by someone who specializes in the student's area of need, often a special education teacher. Because of the multi-tiered approach to education and the various reforms that emphasize standards for practice, evidence-based instruction, and progress monitoring, both general and special education teachers must work as partners to meet the educational needs of all students. It is important for both general and special education pre-service teachers to have the knowledge and pedagogical skills to create effective partnerships and implement multi-tiered quality mathematics instruction for all students.

The implementation of these reforms and RTI are affected by the attitudes and beliefs of the teachers who implement them (Tschannen-Moran & Woolfolk, 2001). To explore this further, researchers have investigated elementary pre-service teachers' efficacy beliefs related to mathematics instruction. However, federal mandates related to student achievement include the progress of all students, including students with disabilities and require the special education teacher to be highly qualified in content knowledge. Thus it is important that research related to mathematics education examine special education teachers' efficacy in mathematics instruction. Specifically, special education teachers must be prepared in providing interventions for students at-risk for mathematics failure, and partnering with general education teachers in the implementation of tier one instruction. For students in tier two and tier three who need effective interventions in mathematical content, it is imperative that their teachers have adequate preparation to provide mathematics instruction. In addition, each school must show students with disabilities are making adequate yearly progress in the grade level curriculum required by the states. Therefore, it is important that special education teachers demonstrate competence in mathematical content as well as the attitudes and efficacy that are conducive to effective teaching for students who struggle in mathematics. There is a paucity of literature regarding mathematics teaching efficacy within the field of special education. However, the existing literature regarding mathematics efficacy of pre-service general education teachers can be used as a framework for exploration of the efficacy beliefs of special education teachers (Charalambous, Philippou, & Kyriakides, 2008; Gresham, 2009; Swars, Daane, & Giesen, 2006).

It is important to review the research related to pre-service teachers' mathematic teaching efficacy to gain insight that can improve teacher preparation and the quality of instruction for all students. Charalambous, Philippou, and Kyriakides (2008) studied the effect of field work on

mathematics teaching efficacy. Eighty-nine pre-service general education teachers completing field experience in disciplines across grade levels and certification areas participated in the study. The researchers found that pre-service teachers' mathematics teaching efficacy changed during their field experience. Daily experiences teaching mathematics and interactions with mentors, tutors, peers, and children greatly influenced these changes. Mentor teachers had the greatest influence on pre-service teachers' mathematics teaching efficacy. The feedback from a well regarded mentor had the greatest impact on the pre-service teachers' efficacy beliefs. Additionally, a large discrepancy between the mentor's teaching style and beliefs and those of the pre-service teacher had a negative influence on mathematics teaching efficacy.

Although experiences during one's teacher preparation program may influence efficacy, perhaps pre-service teachers have certain traits or characteristics that influence their efficacy as well. Swars, Daane, and Giesen (2006) investigated the influence of mathematics anxiety on pre service teachers' mathematics teaching efficacy. Twenty-eight pre-service teachers enrolled in a mathematics methods course participated in the study. The participants completed the Mathematics Anxiety Rating Scale (Richardson & Suinn, 1972) and the Mathematics Teaching Efficacy Beliefs Scale (Enochs, Smith, & Huinker, 2000). Four participants participated in semi-structured interviews based on their high or low mathematics anxiety ratings. The rating scales and interviews showed that high anxiety was related to low mathematics efficacy.

Gresham (2009) extended Swars, Daane, and Giesen's (2006) research by surveying a larger group of pre-service elementary general education teachers who were enrolled in a mathematics methods course. The participants completed the Mathematics Anxiety Rating Scale (Richardson & Suinn, 1972) and the Mathematics Teaching Efficacy Beliefs Instrument (Enochs, Smith, & Huinker, 2000). A portion of the participants were interviewed. The researchers found that pre-service teachers who had lower mathematics anxiety had higher mathematics teaching efficacy. Both the quantitative and qualitative measures suggested that there was a negative relationship between mathematics anxiety and teaching efficacy beliefs.

Bates and Latham (2011) continued the line of research regarding pre-service teachers' mathematics teaching efficacy by investigating its relation to their mathematics knowledge. Eighty-nine early childhood pre-service teachers completed the Mathematics Self-Efficacy Scale, the Mathematics Teaching Efficacy Beliefs Instrument (Enochs, Smith, & Huinker, 2000) as well as the Illinois Certification Testing System Basic Skills Test. The researchers found the pre-service teachers who scored high on the basic skills test rated both their mathematics efficacy and their teaching efficacy higher than those achieved lower scores on the basic skills test.

Research has shown that teachers' mathematics efficacy is influenced by the amount of mathematics coursework (Chang 2009; Swacjamer et al., 2009), mentorship during pre-service field experiences, mathematics anxiety (Gresham, 2009; Swars, Daane, & Giesen, 2006), as well as level of mathematics knowledge (Bates & Latham, 2011. While the aforementioned studies provide insights pertaining to pre-service general educators' mathematics efficacy, they are missing information regarding the mathematics teaching efficacy of pre-service special education teachers. In addition, it is unknown whether there are differences between teaching efficacy of special education and general education teachers. This is significant because special education

teachers at the elementary level provide mathematics instruction through their status as highly qualified teachers per the *No child Left Behind Act* (2002). The same concerns regarding the relationship between teaching efficacy and quality instruction should apply to all elementary-level teachers. The purpose of this study, then, is to investigate the mathematics teaching efficacy of elementary general education and special education pre-service teachers. In addition, this study seeks to extend the literature by investigating general and special education teachers' elementary-level mathematics content knowledge, exploring the relationship between the teaching efficacy and mathematics skills.

Method

Participants

The participants in this study consisted of 178 pre-service graduate and undergraduate students enrolled in either an elementary general or special education program. All participants were seeking initial elementary certification (kindergarten through grade six) and were chosen because the mathematics knowledge surveyed included content through the sixth grade level. The participants who were graduate students had undergraduate degrees in fields other than education and were seeking initial certification in special or elementary education. These students' course of study included the same course content requirements as the undergraduate participants, including coursework in elementary mathematics. The participants were enrolled in two public universities that were equivalent in size, one located within the Southwestern region and the other located in the Southeastern region of the United States. In addition, the teacher preparation programs within each university were equivalent in size. None of the participants were or had been employed as teachers; their only teaching experience involved field experiences within their preparation programs. The participants had completed field experiences in classrooms, but had not completed their internship or student teaching, meaning that none of the participants had been solely responsible for classroom instruction; they had observed and taught under the supervision of a cooperating teacher in the public schools. All of the participants had completed their programs' mathematics content requirements as well as methods courses in teaching mathematics to elementary level students. The general education participants' mathematics methods course had a field component in which pre-service teachers observed and taught in elementary mathematics classes. The special education participants' methods course did not involve a mathematics field experience component as part of the course; the participants' concurrent field experience was in a special education setting in which mathematics may have been one of the instructional areas. Of the students participating in the study, 64% (n = 113) identified themselves as future general educators, while 36% (n = 65) identified themselves as future special educators. Within teacher preparation programs at each university, the size of special education programs as compared to general education programs is equivalent or smaller than the proportion within this study. In addition, the proportion of elementary general education teachers (1,655,800) to special education teachers (459,600) in the United States is slightly less than 3:1, according to the Bureau of Labor and Statistics (2012). Gender demographics were 6% (n=11) male, 94% (n=167) female. Demographic information associated with cultural

background were 5% (n=9) African American, 16% (n=28) Latina/Latino, 76% (n=136) White, 2% (n=3) Asian, and 1% (n=2) Other. The participants' ages fell into the following categories: 18-20 years (20%, n=36), 21-29 years (59%, n=105), 30-39 years (12%, n=21), 40-49 years (7%, n=12), and 50-59 years (2%, n=4). The demographic data are summarized in Table 1.

All of the participants in this study were seeking teacher certification at the elementary level within their respective states. The pre-service special education teachers were seeking an additional certification in special education. Therefore, the two groups were comparable since both groups would be considered highly qualified to teach mathematics to students at the elementary level after completion of their programs.

Survey Instruments

Computational knowledge for this study was surveyed using the *Math Operations Test-Revised* (MOT-R)(Fuchs, Fuchs, Hamlett, & Stecker, 1991). The MOT-R measures mathematical operations skills through the sixth grade level. The MOT-R is correlated (r = .78) with the computation sub-test of the *Stanford Achievement Test* (Fuchs et al.). This instrument was chosen based on the number of items related to each skill. Rather than one item per skill, the participants had multiple opportunities to demonstrate each computational skill.

Mathematical problem solving skills were surveyed using the *Math Concepts and Applications Test* (MCAT) (Fuchs et al., 1994). The MCAT measures mathematical reasoning through the sixth grade level. The items survey knowledge of number concepts, numeration, applied computation, geometry, measurement, charts and graphs, and word problems. The criterion validity of the MCAT with the Concepts of Number subtest of the *Stanford Achievement Test* was .80 and the internal consistency reliability was .92 (Fuchs et al.). This test was chosen based on the variety of skills assessed and the format of the instrument.

The survey packet also included a questionnaire, eliciting demographic information and a mathematics teaching efficacy scale. Participants were asked to identify the following: their (a) age; (b) cultural background; and (c) area of future certification. The participants also completed the Mathematics Teaching Efficacy Beliefs Instrument (MTEBI) (Enochs, Smith, & Huinker, 2000). The MTEBI consists of twenty-one items, thirteen on the Personal Mathematics Teaching Efficacy subscale and eight on the Mathematics Teaching Outcomes Expectancy subscale (Enochs, Smith, & Huinker). The Personal Mathematics Teaching Efficacy subscale relates to pre-service teachers' beliefs in their individual capabilities to be effective mathematics teachers. The Mathematics Teaching Outcomes Expectancy subscale relates to pre-service teachers' beliefs that effective teaching can bring about student learning regardless of external factors. The MTEBI uses a Likert scale with five response categories: strongly agree, agree, uncertain, disagree, strongly disagree. Possible scores range from 13-65 on the Personal Mathematics Teaching Efficacy subscale. Possible scores on the Mathematics Teaching Outcomes Expectancy subscale range from 8-40. Higher scores are indicative of stronger efficacy beliefs. Reliability analysis produced an alpha coefficient of 0.88 for the Personal Mathematics Teaching Efficacy subscale and 0.75 for the Mathematics Teaching Outcomes

Expectancy subscale (Enochs, Smith, & Huinker). Confirmatory factor analysis indicated that the two subscales are independent, adding to the construct validity (Enochs, Smith, & Huinker).

Procedures

The surveys and questionnaire were distributed and completed by graduate and undergraduate students enrolled in general education and special education courses specific to methods within each major. The participants volunteered for the study and completed the background questionnaire at the beginning of a class meeting. The background questionnaire and MTEBI were completed first so that the mathematics tasks within the problem solving and computation survey did not interfere with the participants' METBE ratings. At the next class meeting, participants completed the computation and problem solving surveys using pencil and paper. No time limit was assigned, but surveys were completed in an average of 30 minutes. The order of the computation and problem solving surveys were counterbalanced so that half of the participants completed computation first and the other half completed problem solving first.

Data Analysis and Results

A Multivariate Analysis of Variance (MANOVA) was conducted using the Statistical Package for the Social Sciences (SPSS) version 20. A MANOVA was chosen for this analysis because there were multiple dependent variables and a MANOVA simultaneously tests two or more related dependent variables while controlling for the correlations among the dependent variables (Mertler & Vannatta, 2005). The independent variable was certification, general or special education certification. The dependent variables were: (a) the percent correct scores in computational, (b) percent correct scores in problem solving, (c) ratings on the Personal Mathematics Teaching Efficacy (PTME) Scale, and (d) ratings on the Mathematics Teaching Outcomes Expectancy (MTOE) Scale.

The results of the MANOVA indicated significant differences for certification area, Wilk's $\Lambda = 0.937$, $F(4, 171) = 2.9$, $p < .05$. Univariate analysis indicated that Mathematics Teaching Outcome Expectancy and problem solving performance differed based on certification. ANOVA results for MTOE were $F\ (1, 174) = 4.66$, $p<0.05$, and for problem solving were $F\ (1, 174) =4.9$, $p<0.05$. General education pre-service teachers indicated higher outcome expectancy, whereas special education pre-service teachers had higher problem solving scores. There was no significant effect between general education and special education with regard to Personal Mathematics Teaching Efficacy or computation performance. The means and standard deviations are summarized in Table 2.

An additional analysis was conducted to examine the relationship for both general education and special education pre-service teachers' level of efficacy and their mathematics skills. The pre-service teachers' teaching efficacy scores were divided into three groups (PTME of 38 or less, PTME of 39-51, and PTME of 52 or more) and their outcome expectancy scores were divided into three groups (MTOE of 24 or less, MTOE between 25 and 31, and MTOE of 32 or more). A multivariate analysis of variance was conducted to determine differences among computation and problem solving skills as related to level of PTME and level of MTOE. Differences were

found for PTME, Wilk's $\Lambda = 0.927$, F (4, 336) = 3.23, $p<0.05$. ANOVA results indicate that there are differences in computation and problem solving skills based on all participants' PTME scores with computation F (2, 169) = 4.22, p<).05 and problem solving F (2, 169) = 4.08, p<0.05. Post Hoc analyses show that participants with lower scores (PTME less than 38) demonstrated lower computation and problem solving scores (Tukey HSD, $p < .05$). Therefore, both special and general education pre-service teachers who indicated a lower level of perceived teaching efficacy had lower scores for computation and problem solving skills compared to pre-service teachers who indicated higher levels of perceived teaching efficacy. There were no significant differences in computation or problem solving scores based on general and special education pre-service teachers' perceived level of mathematic teaching outcome expectancies. The means and standard deviations for calculation and problem solving based on efficacy are summarized in Table 3.

Discussion

The purpose of this study was to extend previous research regarding mathematics teaching efficacy to include special education teachers and investigate differences in pre-service teachers' content knowledge or skills. The inclusion of special education teachers into this line of research is needed since special education teachers' responsibilities have changed over the past decade with increased expectations for achievement for students with disabilities and changes in intervention models (e.g., RTI) for students at risk for failure. The participants in this study demonstrated their knowledge of mathematics computation and problem solving skills within content ranging from the kindergarten level to the sixth grade level. In addition, the participants completed the MTEBI, rating their mathematics teaching efficacy and outcomes expectancy.

Mathematics Efficacy and Outcome Expectancy

This study extended the research of Swars, Daane, and Giesen (2006) and Gresham (2009) by including pre-service elementary special education teachers. The participants in the current study rated their teaching efficacy on a scale from 13 to 65 (mean=51.2) and outcome expectancy with a scale from 8 to 40 (mean=29.5) similarly to the ratings obtained by Swars, Daane, and Giesen (teaching efficacy mean=48.9 and outcome expectancy mean=29.1) and Gresham (teaching efficacy mean=50.8 and outcome expectancy mean=31.4). In the current study, there was no statistically significant difference in general and special education teachers' mathematics teaching efficacy. Both groups of teachers reported similar levels of teaching efficacy (51.8 by general education and 50.0 by special education). There was a statistically significant difference between pre-service teachers' teaching outcome expectancy, general education being higher (mean=30. 0) than special education (28.7). However, the difference in scores was less than two points and it is not clear that this is a socially valid difference. This may indicate that there is a slight difference between general education and special education teachers' beliefs that their students will be successful when provided with effective mathematics instruction. The pre-service special education teachers who participated in this study were completing a generic program in which they were prepared to teach children with high incidence as well as low incidence disabilities who participate in the general education curriculum to varying degrees. The slight difference may indicate that pre-service special education teachers

believe that factors other than effective instruction may influence their students' mathematics performance. If statistical difference truly reflects a real difference, pre-service special education teachers may perceive disabilities as factors that influence performance to a somewhat greater extent than effective mathematics instruction. This speculation seems contrary to the beliefs one might expect from pre-service special education teachers. One might expect future special education teachers to advocate for individuals with disabilities and expect that effective and intensive instruction would result in positive student outcomes.

Another explanation of the difference may be program preparation. The coursework related to mathematics in both special education and general education was similar; however, there was a difference in focused fieldwork. General education pre-service teachers completed a focused field experience related to teaching mathematics as part of their methods course and special education pre-service teachers' field experiences are not split by content area. Perhaps experience in an elementary mathematics classroom has an impact on one's teaching outcome expectancy since this is an opportunity to observe outcomes related to teaching methods.

Computation and Problem Solving Performance

The computation skills assessed represented the content that these future teachers will teach to children in schools. There was no statistically significant difference in the mathematics computation performance of general (77% correct) and special education (79% correct) pre-service teachers. There was no difference in participants' performance and this may be a reflection of their preparation. One might expect that general education pre-service teachers' focused field experience might perform differently; however, computation skill may be related more closely with other types of mathematics coursework which did not differ between groups.

Approximately 10% of the participants' computation scores were at or above 90% correct. The computation items that appeared to the be the most difficult, based on errors were: (a) operations such as addition, subtraction, multiplication of fractions and mixed numbers with like denominators (failure to attend to the whole number); (b) adding fractions with unlike denominators (adding both the numerator and denominator or cross multiplying); (c) multiplying decimals (aligning decimals and bringing them straight down to the answer, e.g., 3.25X1.52=949.00); and (e) dividing decimals. Division of decimal numbers appeared to be the most difficult since many responses consisted of series of question marks or comments such as the exclamation, "I don't know!" These computation items represent more complex skills which are similar to the skill areas that are difficult for children, as reported by research related to student achievement (Cawley, Parmar, Foley, Salmon, & Roy, 2001). It is crucial that all elementary teachers can effectively deliver instruction in skills that provide the basis for advanced mathematics study. One might argue that the ability to complete complex computation, such as the division of decimals, would be a pre-requisite for effective instruction of decimals.

Contrary to computation performance, there was a statistically significant difference between general education (80% correct) and special education (84% correct) pre-service teachers' problem solving performance. This finding is contrary to the similarities and differences in preparation between groups described above. Mathematics content coursework is similar and

general education pre-service teachers participate in more focused mathematics field work. Therefore, one might expect this group to perform as well or better than their peers in special education programs. Although special education pre-service teachers performed better than their general education peers, there were similar error patterns across groups. The most difficult items appeared to be: (a) adding standard units of measurement with regrouping (failure to regroup as evidenced by an answer such as 2 yards, 4 feet, and 16 inches); (b) determining the volume of a cube when given the measurements of the height, base length, and base width; and (c) solving a word problem involving multiple steps and multiple operations (inappropriate choice of operations or failure to attend to the need for an operation).

The findings related to mathematics skills show that pre-service special education teachers' mathematics skills are similar to, or slightly better than, those of their general education peers. This is contrary to past criticisms of special education teachers' content knowledge (Brownell, Sindelar, Kiely, & Danielson, 2010). Perhaps these results reflect the changes in special education preparation programs over the past decade, emphasizing content preparation and general education certification at the elementary level. Perhaps the results are reflective of the expansion of content knowledge general educators are responsible for covering as evidenced by the addition of grade levels in some elementary and early childhood state certificates.

It is important to include special education teachers in the research related to effective mathematics instruction and related areas, such as skill and efficacy because they receive certification that qualifies them to teach mathematics content. Due to reforms and RTI, both elementary and special education teachers' responsibilities go beyond collaboration. In a particular, elementary and special education teachers provide interventions to students with and without disabilities through the RTI model. This study is an initial investigation into the skills and efficacy of special education pre-service teachers and it is promising that this group of future teachers demonstrates skill and efficacy similar to their general education peers.

Limitations and Suggestions for Future Research

Results of the current study have limitations. The majority of the sample came from just two geographical regions of the country; therefore, the results may not be representative of the whole country. All of the participants in this study were enrolled in a traditional teacher preparation program, thus the results may not be as realistic because measures were not taken from individuals enrolled in alternative certification programs. In particular, special education is known to struggle with personnel shortages. Due to personnel shortages, the inclusion of pre-service teachers enrolled in alternative certification programs may provide a more realistic measure of competence and efficacy within the field of special education. It is possible that a more inclusive sample of pre-service special education teachers would yield different results. Perhaps individuals within alternative programs would perform differently because their preparation is brief which might lead to lower efficacy (Tissington & Grow, 2007).

Another limitation of this study is its failure to address pre-service teachers' pedagogical knowledge. It is unknown how the participants would have organized instruction related to the

computation and problem solving skills assessed. Furthermore, it is unknown how the participants' content knowledge, teaching efficacy, and pedagogical knowledge are related.

Continued investigation is needed with regard to secondary mathematics content as this study addressed only skills from elementary content. For example, this content could include algebra, geometry, and other areas of mathematics included on high school exit exams. It is not known how general education and special education teachers at the secondary level fare with regard to mathematics efficacy and teaching outcome expectancies. This information would inform current practices, especially with the increased focus on cooperative teaching and requirements of highly qualified status for special education teachers.

Future Research

In order to address some of the limitations of this study, future research might investigate the mathematic skills and efficacy across other regions. Since participants in this study were enrolled in large state universities, future research should include university sites of different sizes and missions. In addition, the inclusion of different types of teacher preparation programs, including alternative certification programs, would also provide a more accurate characteristics.

The current study investigated special education and general education pre-service teachers' skill and efficacy; however, it is not known how pre-service teachers might actually design and implement mathematics instruction. Teachers' instructional practices and instructional interactions with students are a more critical component of improved mathematics achievement of children within our schools. With increased expectations for achievement associated with the *No Child Left Behind Act* (2002), especially for students with disabilities, it is critical that teachers provide effective instruction in the area of mathematics. Future research might explore pre-service teachers' explanations of how they would approach instruction of a particular mathematics task or concept. The inclusion of a qualitative component to this line of research might shed more light on how skill and efficacy relate to instructional practice(s). In addition, the demonstration of evidence-based practices could be investigated with respect to pre-service teachers' skill and efficacy. For example, observations of culminating field experiences could be included in future investigations. Another area of investigation might include the relationship between teachers' mathematics skills and their students' mathematics progress and achievement. Future research might investigate the performance of students as it relates to their teachers' skill and efficacy. Finally, future research could investigate specific skill areas, since the pre-service teachers in this study demonstrated consistent patterns of errors. More emphasis in these areas during preparation programs could address these weaknesses.

References

Bates, A. B., & Latham, N. (2011). Linking pre-service teachers' mathematics self-efficacy and mathematics efficacy to their mathematical performance. *School Science and Mathematics, 111*(17), 325-333.

Brownell, M. T., Sindelar, P. T., Kiely, M. T., & Danielson, L. C. (2010). Quality and preparation: Exposing foundations, constructing a new model. *Exceptional Children, 76*, 357-377.

Bureau of Labor Statistics, U.S. Department of Labor (2012). *Occupational Outlook Handbook, 2012-13 Edition*, Special Education Teachers. Retrieved from: http://www.bls.gov/ooh/education-training-and-library/special-education-teachers.htm

Bureau of Labor Statistics, U.S. Department of Labor (2012). *Occupational Outlook Handbook, 2012-13 Edition*, Kindergarten and Elementary Teachers. Retrieved from: http://www.bls.gov/ooh/education-training-and-library/kindergarten-and-elementary-school-teachers.htm

Cawley, J., Parmar, R., Foley, T. E., Salmon, S., & Roy, S. (2001). Arithmetic performance of students: Implications for standards and programming. *Exceptional Children, 67,* 311-328.

Chang, Y. (2009). A case study of elementary beginning mathematics teachers' efficacy development. *International Journal of Science and Mathematics Education, 8,* 271-297.

Charalambous, C. Y., Philippou, G. N., & Kyriakides, L. (2008). Tracing the development of pre-service teachers' efficacy beliefs in teaching mathematics during fieldwork. *Educational Studies in Mathematics, 67,* 125-142.

Enochs, L. G., Smith, P. L., & Huinker, D. (2000). Establishing factorial validity of the Mathematics Teaching Efficacy Beliefs Instrument. *School Science and Mathematics, 100*, 194-203.

Fuchs, L. S. (2003). Assessing intervention responsiveness: conceptual and technical issues. *Learning Disabilities Research and Practice , 18,* 172-183.

Fuchs, L. S., Fuchs, D. , Hamlett, C. L. ,& Stecker, P. M. (1991). Effects of curriculum-based measurement and consultation of teacher planning and student achievement in mathematics operation. *American Educational Research Journal, 28*, 617-641.

Fuchs, L. S. , Fuchs, D. , Hamlett, C. L. , Thompson, A. , Roberts, P. H., & Kubek, P. (1994). Technical features of a mathematics concepts and applications curriculum-based measurement system. *Diagnostique, 19*(4), 23-29.

Gresham, G. (2009). An Examination of mathematics teacher efficacy and mathematics anxiety in elementary pre-service teachers. *Journal of Classroom Interaction*, *44*(2), 22-38.
Harris-Murri, N., King, K., & Rostenburg, D. (2006). Reducing disproportionate minority respresentation in special education programs for students with emotional disturbances: toward a culturally responsive response to intervention model. *Education and Treatment of Children , 29*, 779-799.

Hoover, J. J. & Patton, J. R. (2008). The role of special educators in a multitiered instructional system. *Intervention in School & Clinic*, 43, 195-203.

Houchins, D. E., Shippen, M. E. , & Flores, M. M. (2010). Math assessment and instruction for students at-risk. In R. Colarusso, & C. O'Rourke (Eds.), *Special education for all teachers*. Dubuke, IA: Kendall/Hunt.

Individuals with Disabilities Education Improvement Acts of 2004, Pub. L. No. 108-446, 118 Stat. 2647 (2004) (amending 20 U. S. C. §§ 1440 et seq.).

Mertler, & Vannatta (2005). Advanced and Multivariate Statistical Methods Practical Application and Interpretation. 3rd edition. Glendale, CA: Pyrczak Publishing

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the national mathematics advisory panel*. Washington, DC: U.S. Department of Education.

Richardson, F. C. , & Suinn, R. M. (1972). The mathematics anxiety scale: Psychometric data. *Journal of Counseling Psychology, 19*, 551-554.

Rivera-Batiz, F. L. (1992). Quantitative literacy and the likelihood of employment amoung young adults in the United States. *The Journal of Human Resources, 27*, 313-328.

Swars, S. L. , Daane, C. J., & Giesen, J. (2006). Mathematics anxiety and mathematics teacher efficacy: What is the relationship in elementary pre-service teachers? *School Science and Mathematics,106*(7), 306-315.

Tissington, L. D., & Grow, A. (2007). Alternative certified teachers and students at risk. *Preventing School Failure, 51*(2), 23-27.

Tschannen-Moran, M., & Woolfolk, Hoy A. (2001). Teacher efficacy: capturing an elusive construct. *Teaching and Teacher Education, 17*, 783–805.

U. S. Department of Education. (2002). *No Child Left Behind: A desktop reference.* Washington, DC: Author.

Table 1. *Participant Personal Demographic Information*

Gender	
Male	11
Female	167
Cultural Background	
African American	9
Latino/Latina	28
White	136
Asian	3
Other	2
Age	
18-20 years	36
21-29 years	105
30-39 years	21
40-49 years	12
50-59 years	4

Table 2.
Means and Standards Deviations Mathematics Efficacy and Knowledge Surveys

Survey	Preservice Teachers	Mean	Standard Deviation
Perceived Teaching Efficacy	General Education	51.84	7.10
	Special Education	50.03	7.94
Outcome Expectancy	General Education	30.01*	3.97
	Special Education	28.68	3.79
Percent Correct Computation	General Education	76.71	10.28
	Special Education	78.79	9.67
Percent Correct Problem Solving	General Education	80.19	11.60
	Special Education	83.90*	8.78

* statistically significant at .05

Table 3.
Means and Standards Deviations of Computation and Problem Solving Scores Based on Efficacy

Efficacy Levels	Computation Mean	Computation Standard Deviation	Problem Solving Mean	Problem Solving Standard Deviation
PMTE scores 52 and Above	77.47	10.56	82.29	10.72
PMTE scores 39 to 51	78.23	9.13	81.62	9.56
PMTE scores 38 and Below	70.17*	11.33	74.50*	15.28
MTOE scores 32 and Above	75.67	12.12	80.03	12.17
MTOE scores 25 to 31	77.77	9.03	81.68	9.86
MTOE scores 24 and Below	79.71	8.77	84.95	9.75

* statistically significant at .05

Accountability for Student Learning: Slow and Steady Progress or Persistent Resistance?

Andrew Q. Morse
State University System of Florida

Abstract

The purpose of this study was to explore the present status of efforts to assess student-learning outcomes within the bachelor's degree granting institutions of the campuses in one system of public higher education. Further, the purpose of this study was also to understand what challenges and criticisms academic leaders report about the call to provide learning outcome evidence. Themes and findings of the study suggest that accountability for student learning is a persistent accountability challenge within higher education institutions.

Accountability is a persistent and commanding policy issue facing higher education leaders, and student learning is a predominant form of accountability evidence requested by stakeholders (National Governor's Association, 1986; Thelin, 2004; Spellings Commission, 2006). Two high level reports provide historical bookends on the persistent call for accountability evidence of student learning. In 1986, the National Governor's Association released a report entitled *Time for Results,* outlining a plan to reform education in the United States. One of the major foci called for nationwide commitment on the part of institutional leaders to focus on improving educational quality and to produce evidence of student learning (National Governor's Association, 1986). Twenty years later, the *Spellings Commission* (2006) released "*A test of leadership: Charting the future of U.S. higher education*." That report once again raised concern with student learning and called for a nationwide effort to re-position the United States as a world leader in education.

Joining these calls have been efforts from lawmakers to push toward institutional accountability for student learning (Flaherty, 2013; Board of Governors, State University System of Florida, 2013). In Iowa, for instance, the legislature passed and the Governor signed a mandate that faculty who teach at each of the three state universities create assessments that provide summative information on student performance and formative results that can be used to generate a plan for improvement (Flaherty, 2013). Florida's state universities are required to develop an '*Academic Learning Compact*' for each undergraduate educational program, which articulates the expected outcomes, corresponding assessment instruments and methods, and a plan for improvement based on the results (State University System of Florida, 2013).

In addition to state-level mandates, accreditation has added pressure to the expectation to assess and report student learning outcomes. Ewell (2001) stated that accreditation organizations, both regional and specialized, have a central role in pushing assessment of student learning outcomes to the forefront at colleges and universities to demonstrate institutional effectiveness. For instance, the Southern Association of Colleges and Schools Commission on Colleges (2012) requires that all educational programs identify a clearly articulated set of student learning outcomes and that such outcomes are accompanied by a plan for how the results will be used to guide improvement. With the

requirement that federal student aid is tied to regional accreditation, universities have an added incentive to prioritize the assessment of student learning outcomes (United States Department of Education, 2013).

But despite longstanding and widespread interest, accountability for student learning persists as a desired, but largely unaddressed, form of accountability evidence. Consequently, this paper seeks to describe the current status and challenges associated with institutional efforts to respond to expectations for learning evidence. This description will not be given, however, without first considering the underlying contextual factors that influence these efforts. Several questions are warranted to give consideration to this context.

Why are stakeholders concerned with accountability in student learning? Many of the students who graduate from colleges and universities are perceived to lack core competencies associated with an undergraduate education. A recent book entitled *Academically Adrift* by Arum and Roksa (2011) criticized colleges and universities for not asking enough of students, arguing that insufficient rigor has led to student incompetence with basic skills such as written communication and critical thinking. Examining test scores, survey responses, and transcripts of over 2,000 students, Arum and Roksa found that students do not significantly improve in these skill domains as a result of the undergraduate curriculum.

In addition, employers commonly report that graduates lack basic skills to meet job-related demands. In 2010, the American Association of Colleges and Universities released a report on employer perceptions of college learning. Employers were interviewed about the skills needed to address challenges associated with a globalizing marketplace and about their perspectives on the effectiveness of institutions to prepare students with the skill sets to meet these challenges. Participants shared that a broad range of skills and a specialized understanding of a field were necessary for success in today's workplaces. However, employers also reported that many graduates entered the workplace insufficiently prepared to meet these expectations and that college and university leaders should address these skill deficiencies (American Association of Colleges & Universities, 2010).

Aside from perceived deficits in the competencies of college graduates, many of higher education's key stakeholders view institutions as incapable of or disinterested in sufficiently addressing accountability and performance expectations (Lane, 2007; Bogue & Hall, 2012). Bogue and Hall (2012) surveyed corporate executives, legislators, and academic leaders in five different states to gather and compare major stakeholder perspectives on higher education accountability. Bogue and Hall found that corporate and legislative leaders perceived that institutions will 'use cosmetic changes to avoid disclosing unflattering information on performance.' Other evidence suggests that resistance to external expectations for change in institutional accountability for learning may insufficiently characterize the issue of institutional response to assessment. Lane (2007) wrote, for instance, that a key factor influencing resistance, where truth to this notion may be observed, can be due to the uncertainty

surrounding the change proposition. What is it, specifically, that stakeholders want as evidence for student learning, it is grounded in consensus, and is there a clear connection between what is expected and what is taught on college and university campuses? At the heart of accountability for student learning may, in part, be the uncertainty about how to navigate tension between an enterprise that is valued for its autonomy in academic mission and evaluated by its duty to demonstrate performance within a democratic social and political fabric. Further, a lack of clarity is present on the extent to which institutions are truly resistant to change or whether this perception is more closely associated with some stakeholder dissatisfaction with institutional mission priorities.

What has been the response from colleges and universities thus far? Despite the complexity of the task to be accountable for student learning, some institutions have taken initiative to provide transparent, comparable data to stakeholders (Association of Public and Land Grant Universities, 2013; The Lumina Foundation, 2013). The Association for Public and Land Grant Universities' Voluntary System of Accountability, a resource in which many colleges and universities elect to participate, launched a pilot program in 2012 that built in evidence of student learning into the profile of key performance criteria on student outcomes and experience indicators. Collected through the pilot project on student learning were scores on standardized exams designed to measure outcomes commonly associated with the general education curriculum (Association of Public and Land Grant Universities, 2012).

Non-profit organizations have also partnered with colleges and universities to address the call for assessment of student learning. For example, the Lumina Foundation (2013) has developed the Degree Qualifications Profile (DQP), which is a resource that seeks to define competencies that a graduate should be expected to know upon completion of an associate's, baccalaureate, or master's degree – though a corresponding, comprehensive plan for how universities can assess and report on such student learning is missing. In partnership with Lumina, the National Institute for Learning Outcomes Assessment, located at the University of Illinois, Urbana-Champaign, has been working to track institutional use of the DQP and found that hundreds of institutions across 46 states are using this framework for assessment (National Institute for Learning Outcomes Assessment, 2013). Through these initiatives, it is evident that many colleges and universities are working to address concerns related to student learning. The impact of these efforts to inform the public and to guide improvement is yet to be determined.

What makes being accountable for student learning difficult? Evidence suggests there is inadequate consensus about what forms of learning evidence are considered desirable among major stakeholders (Bogue & Hall, 2012). In Bogue and Hall's five-state study on corporate, political, and academic leader views on accountability, for instance, there was wide disagreement between the three groups on the desirability of various student learning outcomes as an indicator of performance. Business and state political stakeholders differed from college and university presidents, provosts, and faculty on the desirability of student knowledge of democratic culture and heritage as an indicator of performance. The ability to pursue knowledge through different modes of inquiry was

also viewed as less desirable among business and political stakeholders as compared to academic leaders. The challenge to identify what should be measured may provide difficulty for college and university leaders to effectively demonstrate accountability to stakeholders whose viewpoints vary on the desirable outcomes expected of graduates.

A study conducted by the National Association for College Admissions Counseling (2008) concluded that one-third of students entering a two- or four-year college or university in the United States will transfer to at least one other institution during their time in college. Given overlap in broad competency areas taught at each institution to which mobile learners attend, it may be difficult to target the effectiveness of learning interventions specific to one institution or another. In addition, college is a place where students are presented with multiple opportunities to develop essential competencies across the curriculum. It can also be difficult to ascertain with confidence the extent to which assessment results offer a true reflection of students' learning performance in relation to a particular institution or curriculum.

In the National Center for Public Policy and Higher Education's (2008) *Measuring Up* report, each state was given an A through F grade to indicate performance in an easily digestible manner on a variety of indicators tied to institutional quality and effectiveness. Included in this report was an indicator on student learning; each state earning a grade of "Incomplete" due to a lack of commonly accepted student learning outcomes assessments across the nation. In the report, Peter Ewell noted that since the release of *Measuring Up's* first edition in 2000, fewer institutions seem to be committing to national efforts to provide comparable, transparent information on learning (National Center for Public Policy and Higher Education, 2008). This lack of cooperation, however, is not due to an insufficient array of available instruments to assess student learning. Stanford University's National Center for Postsecondary Improvement houses an electronic inventory of dozens of instruments commonly used to measure college-level learning (National Center for Postsecondary Improvement, 2013). What, then, is stalling progress on assessment as an accountability indicator and a quality assurance mechanism?

The ability for college and university leaders to respond to stakeholder calls for evidence of student learning is met with significant contextual considerations. To understand how college and university leaders are attempting to respond, if at all, to expectations for student learning outcomes involves a deeper look into the steps currently underway to collect this evidence. Given the longstanding concern with accountability for student learning, however, it is evident that significant barriers impede the ability for or willingness of many institutional leaders to gather, report, and use information on student learning.

The Research Problem and Purpose

The problem is that although the call to assess student learning persists as an accountability challenge, there is an insufficient understanding of what steps, if any, colleges and universities are taking to gather such evidence. Further, little is known about the challenges faced by college and university leaders to address stakeholder expectations

for evidence of student learning. This knowledge is important because it adds narrative to institutional efforts to respond to critique and concern expressed by key stakeholder groups expressed through several decades. Consequently, the purpose of this study was to explore the present status of efforts to assess student-learning outcomes within the bachelor's degree granting institutions in one system of public higher education. Further, the purpose of this study was to understand what challenges and criticisms, if any, are reported by academic leaders about the effort to produce evidence of student learning. The study was guided by the following research questions:

> -What efforts, if any, are institutions presently taking to assess and report student-learning outcomes and why?
> -What types of learning outcomes, if any, are institutions trying to measure?
> -What challenges and criticisms, if any, currently impede institutions' abilities to gather learning outcome data?

Method

The study utilized a multi-site case study design. Semi-structured interviews, field notes, and site documents provided information about learning outcome focused initiatives on the three baccalaureate degree granting campuses of one system of public higher education.

Research Sites and Population

University A is a large public research- intensive institution offering bachelor's, master's, doctoral, and professional degrees. Approximately 21,000 undergraduate students attend the institution. Over 300 undergraduate degree programs are offered. University B is a mid-size public institution and offers bachelor's, master's, and doctoral degrees. Approximately 11,400 students attend the institution and nearly 140 undergraduate majors are offered. University C is a small to mid-size public, undergraduate and graduate-degree granting campus. University C enrolls around 7,500 undergraduate students, and offers bachelor's, Master's, and doctoral degrees. Universities A and B have been in operation for over a century while University C was founded in the early 20^{th} century.

The population from which the researcher sought participation included not only the institutional administrators who held official responsibility with learning outcomes assessment, but also those who were responsible for accountability to stakeholders. As such, the presidents or chancellors and chief academic officers or provosts were eligible participants for the study. The researcher then used chaining from these participants to connect with participants who were principally responsible for outcomes assessment on each campus. The additional personnel were the directors of institutional research, college deans, and vice provosts for undergraduate programs. Twelve participants (4 at each campus) were interviewed.

Sources of Data

Three sources of data were used to conduct the study: semi-structured interviews, site documents, and field notes. Site documents included institutional accreditation reports, accountability documents, strategic plans, and others provided to the researcher by the participants. Field notes were taken to provide documentation of observations that further illustrated the participants' responses to the interview protocol that could not otherwise be recorded. The interview protocol was comprised of three major inquiries: assessment instruments, initiatives, and challenges.

Procedure

Upon IRB approval, the researcher contacted eligible university administrators for participation in the study. Though chief executives served as the lead off participants, the researcher was linked to other administrators who held formal responsibility with outcomes assessment. In addition, the researcher was connected to academic administrators who work closely with faculty and college deans with regard to outcomes assessment. In addition to interview data, relevant site documents were obtained through the university website and were also gathered upon recommendation of interview participants. Field notes were also taken to provide additional thoughts or observations while in the university settings with the participants. Data collection continued until no new evidence was gathered.

Data Analysis

During the initial within-site analysis, the researcher reviewed data and wrote a preliminary list of in vivo codes based on the words and text noted in the data sources for each campus (Bogdan & Biklen, 2007). After reviewing the data several times and modifying the list of initial codes according to similarity, the researcher read through the data to assign codes. Upon completion of coding, the principal investigator generated patterns unique to each institution to compare and contrast findings between the campuses during the between-site analysis. To answer the research questions, the researcher reported themes common between the campuses in relation to the three research questions (assessment initiatives, types, and challenges).

Data Trustworthiness

Member checking and data triangulation were used to ensure trustworthiness of the study's analysis and findings (Guba, 1981; Shenton, 2004; Merriam, 2009; Creswell, 2009). The researcher inquired with participants to verify the accuracy of the interpretations of data collected in the study (Guba, 1981, Merriam, 2009). The researcher also gathered multiple data sources through the study, which were used collectively to affirm the trustworthiness of findings noted in each as an independent source (Shenton, 2004; Creswell, 2009).

Results

An important contextual factor became clear after analyzing the interviews, site documents, and field notes: Student learning outcomes assessment on each campus was driven primarily by compliance with state law and regional accreditation standards. The duty to assess, however, was delegated to particular personnel on the campuses.

Accreditation standards articulated that each educational program must identify, assess, and demonstrate improvement on a clearly articulated set of learner outcomes to show commitment to continuous improvement. State law mandated that each public college and university assess and report scores on assessments of general education and major-specific learning outcomes. At each campus, personnel who managed assessment were getting ready to provide student learning data to the state *and* were in the middle of preparing evidence of compliance with standards for their approaching regional accreditation reaffirmation visits. Embedded in their efforts to comply with these external stakeholders were perspectives on the instruments, initiatives, and challenges associated with student learning outcomes assessment.

Findings Related to Assessment Types

The case study included an inquiry into what learning, specifically, is being assessed at the institutions. Also, the inquiry focused on what instruments are used to produce the assessment results. Two findings related to what is being assessed and by what instrument were exhibited: general education and major-field testing. The state agency's requirements for outcomes assessment required that, with agency approval of the instrument, institutions develop or adopt measures for major and general-education learning that provide a numerical comparison to a standardized or historical score. Each institution adopted a different general education outcomes assessment instrument and a broad array of instruments for major-specific outcomes. Further, some programs were exempt from the state's assessment standards due to a curricular focus that does not yield easily to quantitative assessment (i.e. studio art).

General Education Testing

To comply with state accountability standards for general education assessment, University A used a quantitative instrument to assess critical thinking skills. The instrument is a nationally and internationally standardized instrument and is widely used as a measure of college-level critical thinking skills. Critical thinking ability is delineated in sub-scores with the intended purpose that the results can be used to identify possible curricular improvement areas. Students were tested during the senior year. University B used a value-added assessment to comply with general education assessment standards. Students were tested at the start and completion of the curriculum. The instrument is a widely established, standardized instrument offered in multiple-choice format. A written essay accompanies the assessment and is used to examine writing and analytical reasoning skills. Questions assess students' aptitudes for reading, writing, mathematics reasoning, science-reasoning, and critical thinking. University C offers a

different value-added examination than University B, but students, too, are tested prior to and at the conclusion of the curriculum. This assessment also examines core competencies associated with the general education curriculum as noted in the other instruments. Once every five years, each institution is required to report its student learning outcomes data to the state agency to comply with accountability standards. To incentivize improvement in the results over years, universities are eligible for state performance-based funds if the score from the most recent reporting period improved from the prior period.

Major-Field Testing

Each university developed and adopted a variety of quantitative assessment instruments to assess learning outcomes associated with academic programs. At University A, for example, the investigator collected and reviewed the Department of History's state agency-approved major field assessment. The history format is a multi-question, true-false examination about specific events and individuals noted in the curriculum.

Standardized major field tests, such as those administered for programmatic accreditation, are given to students within departments that seek such accreditation. One example cited in the study was accounting, which at University C is accredited by the Association to Advance Collegiate Schools of Business (AACSB). As part of the accreditation review and reaffirmation process with AACSB, institutions must report Certified Public Accountant (CPA) licensure examination pass rates.

For some academic programs, the departmental faculty members were engaged in embedded assessment of learning outcomes to provide additional information not yielded from the state-mandated quantitative instruments. For instance, Judy, Associate Provost for Undergraduate Learning at University A, shared:

> In my discipline, we teach a capstone course which is intended to bring together all the learning that students have done throughout their studies…they [students] have to present on a media campaign that they have developed. They have to demonstrate that they know the material through integrating their knowledge into the project. As a faculty, we look for student demonstrations of their learning in the presentations. One of the things that we do as a faculty after the presentations is to sit down together and go over our evaluations of the projects and discuss where the students did well, and where they did not do so well.

Though assessment was present at each institution, participants expressed that more work was needed to get each department involved, to show skeptics the utility of assessing outcomes, and to use the data to strengthen teaching and learning. As a result, university-

wide initiatives were underway at each institution to improve upon the assessment process.

Assessment Initiatives

Each campus had built into their institutional strategic plans the priorities to assess learning in response to compliance expectations from the state and from the assessment standards articulated by the regional accrediting body. Therefore, themes noted were *working toward compliance* and *trying to engage in continuous improvement*.

Working toward Compliance

Though institutions were in full compliance with state assessment regulations, meeting these regulations did not translate to fulfillment of the more rigorous standards associated with regional accreditation. State standards only required that scores improved from one reporting period to another, whereas accreditation compliance meant a much more comprehensive approach of documenting, assessing, and planning improvement in student learning. For instance, Gerald, Dean of the College of Arts and Sciences at University B, shared that "some of the departments in my college do not get as much out of the [assessment] experience as others; they aren't that serious about assessment."

In preparation for each university's accreditation reaffirmation visits, the participants noted that, although they will not fully comply with standards articulated by the regional accrediting body, little worry was devoted to what this would mean to the university in terms of their reaffirmation prospects. For instance, Sharon, University C's Associate Provost who was responsible for undergraduate assessment, reported, "Right now, we're coming up on [our regional accreditation reaffirmation] and we're going to get dinged on our assessment. We do a lot of data gathering; we just don't do a lot with it…"

Getting 'dinged' meant that although the accrediting agency may raise concern with the campus' efforts to assess student learning, there was little concern that any serious consequences would come from the finding. Only getting 'dinged,' however, did not mean that the participants were willing to concede progress toward compliance with accreditation.

Though the universities may be unable to fully address the accreditation standard in time for their upcoming reaffirmation visits, participants articulated that moving step by step toward the standard was an ongoing priority. On each site, the participants stated that the immediate first steps are to ensure that each program has a clearly articulated set of outcomes. Then, the campus' assessment leaders plan to have each program develop corresponding instruments and a plan for how the data will be collected. To achieve these goals, the participants were busy integrating assessment into already established quality assurance processes. As Provost at University A, for instance, Allison stated her priorities in relation to getting all of the departments engaged in assessment:

> I've been fine-tuning our academic program review process. We're building student learning outcomes into the program reviews we conduct internally, which our institution requires every five years. Once we do that, our student learning outcomes assessment process will be built into the accreditation process.

At University C, the process involved implementing a new centralized database to provide greater collaboration between departments and university administrators in its established program review process. The database uses a standard structure for departments to enter the learning outcomes, assessment plans, instruments, results, and plans for improvement of academic programs. From there, Lynn, Vice Provost for Undergraduate Programs at the institution states that we "now have a common structure in which to plug the data. Now it's easy to track departments and say, "Well, you don't have anything in there you need."" The ability to make campus-wide steps forward and to provide continuous guidance and support were viewed as critical steps toward fulfilling accreditation standards.

Trying to Engage in Continuous Improvement

Continuous improvement in teaching and learning was also a work in progress. For instance, University A was in the midst of a campus-wide effort to tie learning outcomes assessment data into improvement of the general education curriculum. The participants shared that this process involved gathering faculty members who teach general education courses together to define the competencies associated with the curriculum. That work group is led by Judy, the Vice Provost for Undergraduate Learning, who shared that, "we are right in the middle of the process of making some changes to our general education program for undergraduate students. One of the charges to the committee was to be very intentional about what our learner outcomes are for general education, how we can assess those, and how we can use the assessments to go through an improvement process." In the report released by this work group were some general competencies that students should gain as a result of the general education curriculum. Up to this point, the effort had not tied specific courses to the competencies or provided an overview of how assessment might be used to improve curricula.

On each campus, the ability to design assessment systems was a key area of focus in trying to engage in continuous improvement. But in attempting to move forward, participants shared that two key challenges encountered were to determine which instruments to use and, in turn, to develop a plan based on the type of information yielded by the results. The quantitative instruments used to comply with state assessment standards were met with criticism by faculty in terms of the ability of instruments to yield data that could target specific areas of improvement. For instance, John, University B's Director of Institutional Research, shared that "We get faculty who look at the questionnaires we administer [for compliance] and say, "Well, this doesn't represent

what I teach."" In response, John asks, "Well, what would?" He's also taken leadership on an institution-wide effort to provide answers to this question within each academic program, but not all units have agreed to participate.

At University B, the effort to get programs involved in the continuous improvement goal have thus far resulted in push for embedded assessment, which Susan, the university's Provost described as "a process where faculty take the work already submitted by students in coursework and re-evaluate it for competencies in the curriculum." According to the participants, this type of information provides clarity about what, if anything, should be addressed that a number may be unable to discern with credibility and confidence. The participants also felt that embedded assessment provided a means for faculty to link identified outcomes to specific courses, thereby targeting improvement areas within the curricula.

One illustration of how embedded assessment was cited as a useful resource for strengthening efforts to continuously improve program delivery was in the Department of Music. Lynn, Vice Provost for Undergraduate Programs at University C, reported that:

> Embedded assessment allows us…to look at whether the interpretation of the piece is appropriate for the time period... It's subjective, but the faculty have been able to develop a way to examine student performance because they are qualified in their areas so they can judge.

The example provided shows how universities are adapting to the expectation that assessment lead to continuous improvement, but it also illustrates a divide between the call for comparable quantitative data to evidence learning and the expectation that such data be used to guide improvement. But, from the participants' perspectives, it was a priority for them to make progress by doing what worked to address accountability expectations and strategic goals. However, there was no clear timeline identified for when the universities would achieve their strategic plan goals for institution-wide assessment. To meet this goal, key challenges would need to be addressed.

Challenges to Reaching Assessment Goals

In moving their campuses forward, the study participants encountered challenges that impeded the ability to achieve institutional goals to assess student learning and to use the data to guide improvement. Namely, the themes related to these challenges were resistance based on established practices and concern with assessment decision utility. In turn, before efforts could be made to strengthen assessment systems in response to institutional goals and compliance standards, these participants shared that much time and effort is currently spent addressing criticism, building trust, and gaining support for the value of assessment.

Resistance Based on Established Practices

Participants stated that their efforts to assess and report student learning outcomes were met with resistance from some academic units, particularly where the faculty viewed student learning outcomes as unnecessary or redundant to grades as an indicator of performance, and, as a result, of little importance when it comes to using the data to guide performance. Many of the participants, like Richard, Dean of University C's College of Business, for example, are often asked what value assessment brings that grades do not when attempting to work with programs that do not currently assess outcomes.

In response to the value that assessment data bring to informing the conversation of improving teaching and learning, Allison, Provost at University A, shared, "If I have assessment data, I can say look, a Classics graduate from this institution should be able to do these 10 things, and, doggone it, they can," instead of, "Well, this person got a 3.5 GPA in X number of classes."" Allison's point is that assessment provides an organized crosswalk between what students should be learning and, by linking these outcomes to the courses in which the outcomes are taught, evidence can be generated to target improvements and to demonstrate performance to stakeholders. Grades and grade point averages, in contrast, fail to indicate where students may not be learning expected outcomes of students.

In addition to grades and grade point averages, the participants reported that some colleagues resisted assessment because, not having familiarity with the practice, they worried about its use and intent. Richard, Dean of University C's College of Business, shared that faculty members are often distrustful of how the assessment data will be used. He reported, "It is difficult sometimes convincing faculty members that the data will not be used against any particular faculty member, and, of course, we're not gathering the data in order to get rid of faculty member X, and the convincing process takes a while." In response to this concern in his college, Richard assembled a group of faculty across the departments to serve in a collaborative manner and to act as liaisons back to the faculty to help demonstrate how data are used. In addition, the faculty help communicate to Richard how assessment is being used to improve teaching and learning. As a result, assessment goals were being met within the college and faculty had a clearer understanding of the intent of the assessment process.

Concern with Assessment Decision Utility

Participants encountered and perceived concern with the decision utility of assessment, also described as the ability for the assessments to yield data that could truly measure learning or guide improvement. Bill, University C's Chancellor, expressed concern that the college experience was larger than "a litany of assessment instruments" could gauge. He reflected:

> I'm asked to tell stakeholders all the time about what transpired within our students during college and I don't know. They [our students] started years ago and a lot has happened to them and not just during college. Those who make laws or policy think we produce a product or service – that it can be measured and if we can't show we're doing it then we have to get rid of what we're doing or change it.

Another challenge was that of student motivation to take assessment seriously, particularly where examinations did not directly affect students based on their performance. John, University B's Director of Institutional Research, shared that he often receives completed tests where students fill out the bubbles in various shapes or in a Christmas tree pattern, thereby raising questions about how seriously the instruments can be taken in making judgments about student learning.

Participants shared concern with the ability to draw conclusions about where improvements can be targeted due to broad overlap in the competencies reinforced in the curriculum. At University A, Judy reported, for instance, "We report the data [on general education testing] and show comparisons across colleges, but that's not assessment. That's testing." The distinction is that data should indicate where, specifically, the learning is or is not taking place and should provide useful evidence. The general education examination provides statistical means that indicate performance relative to other colleges within the institution or to other institutions across the country. However, the participants shared that comparable data falls short of being able to guide curricular improvement.

Though participants encountered and recognized the limitations of assessment data, particularly in relation to accountability expectations, they also work to overcome resistance as a means to achieve institutional quality assurance goals for continuous improvement. Those responsible for leading student learning outcomes assessment at each campus had engaged colleagues through initiatives to gather evidence that could yield improvement-oriented results. As a result, many departments recently started using embedded assessment where faculty use assignments from coursework to evaluate whether students learn key competencies.

At the universities, however, efforts to address concerns were a work in progress that still faced criticism and challenges to institution-wide commitment. Not all departments adopted methods that addressed concerns regarding decision utility. Though progress had been made, participants still reported that more work is needed to fully implement assessment processes that responded to institutional goals for quality assurance in teaching and learning.

Discussion

What is stalling progress on outcomes assessment? Consistent with Ewell's (2001) essay on the role of accreditation, regional accreditation was cited as a major factor

driving assessment on each campus. But the participants shared that not complying with assessment standards raised little concern, citing that they are likely to 'get dinged' for not being in full compliance. Without serious reprimand for compliance failure, institutional leaders and curriculum leaders have little motivation to take the practice of outcomes assessment seriously. Will it take the threat or actual loss of regional accreditation to more heavily incentivize learning outcomes assessment? Up to this point, the reinforcement of standards has not yielded satisfactory results at prompting many institutions to prioritize learning outcomes assessment.

Despite the wide variety of instruments used to assess major-specific learning on each campus, and further recognizing the ability for each institution in this study to choose from a litany of general education and major-specific assessment instruments, the issue of decision utility still persisted as a challenge to the institution-wide establishment of assessment practices. Participants from each campus reported concern with the ability of assessment instruments, particularly quantitative measures, to yield results that are useful at guiding curricular improvement or at reflecting performance to stakeholders accurately. Though participants were working to steer the campus forward with assessment, their reports only reflected partial, incomplete engagement across the academic units. Could 'engagement lag' be a major reason why national efforts over several decades have thus far failed to produce widespread commitment on the part of colleges and universities to respond to calls for evidence (National Governors Association, 1986; Spellings Commission, 2006; Ewell, 2008)?

As was mentioned earlier, the presence of student learning outcomes assessment is not a new phenomenon. Yet, in response to accountability standards, why do instruments not viewed as capable of producing useful results still get used? With no shortage of assessment instruments or approaches, it is unclear why college and university assessment leaders would continue to use instruments and approaches that do not yield meaningful, credible results among campus faculty (National Center for Postsecondary Improvement, 2013). Participants were committed to compliance with standards articulated by stakeholders in government and accreditation. But outcomes assessment is not just about compliance or accountability. If a divide exists between what is viewed as credible by stakeholders and what is viewed as useful by faculty, then campus-wide initiatives should extend beyond compliance and focus also on improvement. With regard to general education assessment, this was not observed on each campus, and differing engagement with academic units was reported on devising approaches that responded to criticism about major-specific assessment.

Although the dialogue up to this point has principally focused on learning outcomes assessment in relation to external accountability pressure, this attention is not to suggest that institutions have completely failed to respond to expectations for evidence of student learning and continuous improvement (National Institute for Learning Outcomes Assessment, 2013). To this end, it is also important to consider how stakeholders themselves fit into the picture. Why, despite evidence of progress on student learning outcomes assessment, does there seem to be a lack of recognition by external stakeholders on the efforts of colleges and universities to respond to expectations for

accountability in learning? To what extent do expectations held by state leaders align with those in external accreditation? Perhaps colleges and universities are caught in an accountability paradox where unflattering evidence of performance weighs significantly more heavily than achievements. Added to this paradox is the magnitude of higher education's stakeholder audience whose expectations on performance and mission priorities may often exist without consensus (Bogue & Hall, 2012).

Participants also shared that although they encountered resistance from some faculty and departmental leaders with regard to student learning outcomes assessment, they were also able to make small steps toward achieving institution-wide commitment by engaging colleagues in dialogue about what *would* give them evidence needed to improve teaching and learning. This study suggests that if assessment remains as an accountability priority, and if leaders can demonstrate value to colleagues, perhaps slow and steady progress can be made toward achieving institution-wide commitment to accountability for student learning through the practice of assessment. However, with widespread and persistent public interest in learning outcome evidence, combined with a continued resistance noted at each campus, it is unclear when or if institution-wide learning accountability will be attained.

References

American Association of Colleges and Universities. (2010). *Raising the bar: Employers' views on college learning in the wake of the economic downturn.* Retrieved from http://www.aacu.org/leap/documents/2009_EmployerSurvey.pdf

Association of Public and Land Grant Universities. (2013). *Student learning outcomes pilot project*. Retrieved from http://www.voluntarysystem.org/slo_pilot

Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago, IL: University of Chicago Press.

Board of Governors, State University System of Florida. (2013). Board of Governors regulation 8.016 – Student learning outcomes assessment. Retrieved from http://www.flbog.edu/about/regulations/regulations.php

Bogdan, R., & Biklen, S. (2007). *Qualitative research for education: An introduction to theories and methods (5th edition).* Boston, MA: Allyn & Bacon.

Bogue, E., & Hall, K. (2012). Corporate, political, and academic perspectives on higher education accountability policy. *College and University*, *87*(3), 14-23.

Creswell, J. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles, CA: Sage Publications.

Ewell, P. (2001). Accreditation and student learning outcomes: A proposed point of departure. Retrieved from http://www.chea.org/award/StudentLearningOutcomes2001.pdf

Ewell, P. (2008). Stuck on student learning. *Measuring up 2008: The national report card on higher education.* Retrieved from http://measuringup2008.highereducation.org/states/index.php

Flaherty, C. (2013, July 19). Assessment: It's the law. *INSIDE HIGHER ED*. Retrieved from http://www.insidehighered.com/news/2013/07/19/iowa-state-legislators-mandate-course-level-continuous-improvement-reporting-mixed

Guba, E. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries, *Educational Communication and Technology Journal, 29*, 75–91.

Lane, I. (2007). Change in higher education: Understanding and responding to individual and organizational resistance. *Journal of Veterinary Medical Education, 34*(2), 85-92.

Merriam, S. (2009). *Qualitative research: A guide to design and implementation.* San Francisco, CA: Jossey-Bass Publishers.

National Association for College Admission Counseling. (2008). *Report of the commission on the use of standardized tests in undergraduate admission*. Retrieved from http://www.nacacnet.org/PublicationsResources/Research/Documents/TestingComission_FinalReport.pdf

National Center for Postsecondary Improvement. (2013). *Inventory of higher education assessment instruments.* Retrieved from http://www.stanford.edu/group/ncpi/unspecified/assessment_states/instruments.html

National Center for Public Policy and Higher Education. (2008). *Measuring up 2008: The national report on higher education.* Retrieved from http://measuringup2008.highereducation.org

National Governor's Association. (1986). *Time for results:* A transcript of proceedings of the
National Governor's Association's First Plenary Session. Retrieved from http://www.nga.org/files/live/sites/NGA/files/pdf/1986NGAAnnualMeeting.pdf

National Institute for Learning Outcomes Assessment. (2013). *Degree qualifications profile corner.* Retrieved from http://learningoutcomeassessment.org/DQPCorner.html

Shenton, A. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information, 22,* 63-75.

Southern Association of Colleges and Schools. (2012). The principles of accreditation: Foundations for quality enhancement. Decatur, GA. Retrieved from http://www.sacscoc.org/pdf/2012PrinciplesOfAcreditation.pdf

Spellings Commission. (2006). *A test of leadership: Charting the future of U.S. higher education* (A report of the commission appointed by Secretary of Education Margaret Spellings). Retrieved from http://www.ed.gov/about/bdscomm/list/hiedfuture/reports/pre-pub-report.pdf

The Lumina Foundation for Education. (2013). *The degree qualifications profile*. Retrieved from http://degreeprofile.org/

Thelin, J. (2004). *A history of American higher education.* Baltimore, MD: The Johns Hopkins University Press.

United States Department of Education. (2013). *Accreditation & Participation.* Retrieved from http://www2.ed.gov/admins/finaid/accred/edpicks.jhtml?src=ln

Value Added Methods: Moving from Univariate to Multivariate Criteria

David Newman
Florida Atlantic University

Isadore Newman
Florida International University

Carolyn Ridenour
University of Dayton

Jennifer Morales
Florida International University

Correspondence related to this article should be directed to Dr. David Newman at dnewma14@fau.edu

Abstract

The authors describe five value-added methods (VAM) used in school assessment as the backdrop to their main thesis. Then they review the assumptions underlying measurement and evaluation, the foundation of all assessment systems, including value-added. They discuss the traditional criterion variable used in VAM: a standardized test score. Next, they challenge the univariate assumptions of VAMs, and argue that a multivariate paradigm of VAM is more advantageous for educators and stakeholders. Finally they describe a potential scenario whereby a multivariate VAM might be implemented.

Political pressure for accountability continues to generate support for value added models to measure the effects of public schools (Chetty, Friedman, & Rockoff, 2011; Papay, 2011) both in the USA and across the globe (Timmermans, Doolaard, & de Wolf, 2011). As a result, value added models increasingly generate scrutiny by both researchers and policymakers (Briggs & Weeks, 2011; Glazerman, Loeb, Goldhaber, Staiger, Raudenbush, & Whitehurst, 2010; Jerald, 2009; van de Grift, 2009; Yeh, 2012). Previous initiatives like No Child Left Behind (2001) and now Race to the Top, School Improvement Grants, and the Teacher Incentive Funds, require schools and teachers to meet AYP (Annual Yearly Progress). This was an attempt to quantify the building/district subgroups' success in meeting the current year's goal toward all children being proficient in math and reading by 2014. However, policymakers are beginning to back off that goal. Nineteen states have had the requirements waived, so far, on the condition that each of those states will develop credible alternative assessment plans (Perez-Pena, 2012). Strong alternative plans at the state level might incorporate value added even more enthusiastically as state level bureaucrats attempt to earn the waiver.

It is likely, therefore, that value added models will remain an important form of accountability. At present, measures of outcomes are required by student subgroup (socioeconomic status, ethnicity, special needs). They do not, however, indicate what teacher practices work in generating those outcomes. In other words, AYP does not identify which teachers were effective and which teachers were not effective in contributing to student growth. Teacher effectiveness is directly related to student success and subsequent school / district success. Using student test

scores to assess teacher effectiveness, however, is one of the most controversial of educational evaluations (Caillier, 2010; Hill, Kapitula, & Umland, 2011; Sanders & Horn, 1994).
Many schools and districts have begun to use value added scores in determining teacher effectiveness. Value-added scores (VAS) represent student growth, attributed to a specific time, agent or experience (Chetty, Friedman, & Rockoff, 2011; Gong, 2006). VAS are based on student achievement scores over time. The Department of Education's Race to the Top initiative urges states and districts to uses this summative, high risk, evaluation to identify teachers that they want to hirer, lay off, promote, or give tenure (Goldhaber, 2010). These polices have resulted in districts using results bases accountability systems (Anderson, 2005; Murphy 2012).

However, some schools have used the value added modeling in other more beneficial ways. These schools and districts have moved away from a onetime summative, high risk, evaluation of teachers and started implementing formative, low risk, evaluations. By using bench mark scores throughout the year schools have been able to identify which practices (time, agent and/or experience) are effective and which ones are not. Student assessment (and consequently student growth) can be used as a tool for educational improvement. By identifying where teachers are weak, professional development can be designed to target specific teacher needs. Combining the professional development with the best practices identified by the data, schools are better able to implement these practices for improving student growth and thus leading to more effective schools (Burnett, Cushing, Bivona, 2012).

Calculation of VAS seems straightforward because they are based on student growth; but, there is not one standardized value-added model (VAM) that policy makers, educators, and educational researchers can agree upon. In fact the models vary widely. Individual models attempt to correct for the weaknesses of other models. All strive to use student growth to calculate teacher/school/district effectiveness. The challenge is to attempt to control for all possible variables that covary with teacher instruction. One of the major criticisms of VAMs is their unidimensional limitation. In other words, none of the VAMs are multi-dimensional, i.e., incorporate more than one outcome measure: change in student test scores.

Purpose

The purpose of this paper is to (1) discuss five value-added methods (VAM) used in school assessment, (2) review the assumptions underlying measurement and evaluation, (3) discuss options for the criterion variable and how the criterion variable is selected, (4) challenge the univariate assumptions of VAMs, and (5) argue that a multivariate paradigm of VAM is more consistent with the principles of good measurement, more helpful to accountability purposes, and more likely to increase the meaning of results for educators. Regarding this fifth purpose, we extend one of the conclusions of Timmerman et al. (2011) in this journal, i.e., that the Netherlands school system, among others, might include multiple outcomes that go beyond only cognitive measures. In our discussion, we also propose ways that a school district might implement such a plan.

Value-Added Methods in School Evaluation: Five Models

In *A Review of Value-Added Models* Hibpshman (2004) reviewed four of the most common value added models for the Kentucky Education Professional Standards Board. These four basic models are discussed next.

In Hibpshman's (2004) review, the first of the four most common models is the fixed effects model (FEM) where teachers, classes, and schools are treated as fixed effects. Fixed effects means that the variable's error term is based on the assumption that the independent variable has no variability and that it is not a sample of a larger population. The advantage of this model is simplicity. One can think of this as a simple fixed effects analysis of variance model, where the dependent variable is the state test and the independent variable could be teachers, classes or schools. This VAM tends to answer research questions pertaining to significant growth of students that can be accounted for by teachers or schools. The FEM does not assume that the teachers or school are a sample of a population of teachers or schools, but, instead, assumes that the growth is for these specific teachers and schools.

The second model is a simple fixed effects model (SFEM). It is different from the FEM. In this model, the effect size of one building in a district is analyzed and compared to another building in the same district. The SFEM model does not employ data on confounding factors and is intuitive in nature. The research question answered by this model focuses on the effect from differences between schools without taking into account differences in teachers or students.

The third model is the layered mixed effects model (LMEM), the model used by Sanders and Horn (1994, 1998) in the TVAAS. The LMEM uses student change scores with randomized school effect. A randomized effect means that the variable's error term is based on the assumption that the independent variable is a sample from a population. This model assumes that it accounts for confounding variables because these variables are actually nested within each student and are therefore controlled for by multiple measurements of each student. The research question that this type of model answers is very similar to the FEM except that it treats schools as random. The LMEM also looks at layers, such as a school layer or a classroom layer.

The fourth VAM structure is the hierarchical linear model (HLM), which also assumes a random school effect. The HLM allows for one to control for covariates at both the school level and the student level. In addition, HLM also tends to assume random effects at the district level, teacher level, and student level. Like the LMEM, the HLM answers the question of school or teacher effect on student growth, but instead of treating only the school as random it can also treat teachers and students as random. This model allows the researchers to examine and control the relationship between nested groups. For instance, students are nested within classrooms, teachers are nested within schools, and schools are nested within districts.

A fifth type of VAM, constructed by The Reading First Ohio Center (RFOC), is an addition to those in Hibpshman's (2004) review. Including that here adds to the background discussion on VAMs. Similar to a fixed-effects model as discussed above, but, instead of focusing on the effects at teacher level it focuses on the effects at a program level.

RFOC was interested in the programmatic effects that were results of the professional development and financial support provided by RFOC. For the RFOC model, the students were

separated into three groups: At Risk, Some Risk and Low Risk groups. These designations were assigned to students based on their most recent Dynamic Indicators of Basic Early Literacy Skills scores (DIBELS) (Good & Kaminski, 2002). Growth was simply calculated as the difference between the two most recent DIBELS scores. To get the clearest picture of how effective the program was, this information was disaggregated by grade, by schools, by district typology (urban, suburban, and rural) and by race. This resulted in the following graphs.

Figure 1 shows the typical shed pattern, indicating that the At Risk group has made the largest gain in reading scores. This would suggest that the RFOC money focusing on interventions for the At Risk group was well spent. Figure 2 shows a typical tepee pattern. In this figure the greatest gain in DIBELS score is with the Some Risk students. The At Risk and Low Risk students show little gain. This indicates that the teachers are teaching the curriculum but may not be differentiating based on student needs, or that the intervention support might not have been structured effectively. Figure 3 shows an upward shed pattern, the reverse of the pattern shown for the At-Risk group.

Assumptions Underlying Measurement and Evaluation

\The five value added models provide some background that provides a common understanding on which we can begin to argue for moving VAM from its limited unidimensional structure to a more beneficial multi-dimensional structure. Before making that argument, however, we need to lay out our perspectives on measurement, testing, and validity to complete that foundational understanding.

Test validity refers to the "meaningfulness and appropriateness of the uses and interpretation…of assessment results" (Linn & Miller, 2005, p. 100). Tests are not valid in isolation; they are valid for particular purposes. To assess student learning or growth with paper/pencil tests, one needs an entire battery of tests, not a single measure. At least three arguments support this need. First, one test has insufficient validity to answer the question about a student's academic performance. One "principle of measurement validation," according to Nitko (2004) is that only "after combining several types of evidence" can the researcher judge it in relation to some intended use (p. 56). Second, a core tenet in measurement is the content representativeness of the domain being measured (Nitko, 2004). In other words, is whatever is being measured representative of the performance domain? For educational accountability purposes, the entire domain of what's being measured must be considered. Third, standardized tests that are used for assessing school effectiveness are valid for drawing group conclusions but not valid for drawing conclusions about individuals. For individual students, standardized tests may be effective screening devices with scores becoming the basis for hypotheses about a student's achievement. The hypotheses are pursued by administering other tests to confirm or disconfirm the results of the group tests.

A gap separates valid measurement practice and the reality of contemporary testing programs in schools. While providing a battery of tests, rather than a single test, is necessary from a psychometric perspective, it is rarely if ever implemented in schools, whose public purposes are sometimes antithetical to good measurement. Standardized test scores become measures of student performance for both group and individual student purposes. Parents receive summary documents about their children's academic performance on standardized tests. There is no

similar standardized test score that represents other types of social or emotional learning that the student has experienced over a year. School reform efforts over at least the past two decades that have resulted in state and federal laws mandating that schools be accountable for the learning of all children (despite demographic differences); however, the academic achievement test is the only required measure of that learning. Therefore, schools have no incentives to use batteries of tests. The subject matter test score stands.

For example, if a student fails the Ohio Achievement Assessments (OAA) in math but has a record of earning the highest grades in all math classes at his/her home school, that discrepancy is not explored when reporting OAA results. No further study of the discrepancy between these two measures of student's performance is carried out. Likewise, a student who passes the same OAA in math with flying colors yet earns Ds and Fs in his/her home school math classes is unlikely to be studied further. Parents, guidance counselors and teachers cannot explain the differences without further analysis. However, for accountability purposes, the standardized test score stands.

The Criterion Variable in Value Added Models

The VAMs reviewed so far use a test score, or the difference between two test scores, as the dependent or criterion variable in calculating the value-added score. Most VAMs include some combination of school factors and/or non-school factors as independent variables. School factors are those that schools can control, such as class size or teacher salary. Non-school factors are those that schools cannot control, such as ethnicity or socioeconomic status of the students. How valid is the single test score, the criterion variable, in representing a student's growth? To reiterate, standardized test scores are the sole measure of student achievement.

Standardized test scores do not capture growth in other school or non-school areas such as self-efficacy, organizational skills, critical thinking, and emotional skills. However, schools target such skills in their mission statements and strategic plans. Buildings implement programs to address these school and non-school domains of learning. This being the case, researchers admit that adequate assessment of student growth requires a battery of tests (not just one test). As a result, researchers are necessarily led to multivariate analysis to quantify student growth. None of the VAMs discussed so far use multivariate analysis.

Considering Multiple Criterion Variables in a Value Added Model: One Example

Statistical analysis is a way of partitioning variance to look at it more completely, to better understand the variability in the dependent variable. The problem with all of the VAMs is that they are unidimensional. Every VAM derives its impact or its effect as measured against one specific test score. The test might be the Ohio Achievement Assessments, or in the case of RFOC, the DIBELS, but all VAMs are based on scores from one test. This limits the scope of the generalizability of these models. A multivariate technique might provide a more comprehensive and accurate model. For instance, if one includes other factors such as grade point average, a measure of portfolio presentation, emotional stability, and skills in critical thinking, there would no longer be one factor (test score) but as many as four or more factors that could then be utilized to create a much more comprehensive and representative value-added

score. Burnett et al. (2012) stated that since teaching by its nature is multifaceted and therefore a multiple measure approach would better capture the teacher's true effectiveness. The Measures of Effective Teaching project found that it was a combination of teacher observations, student feedback, and the VAS that provided the indicator of teacher effectiveness (Kane & Staiger, 2012).

This multivariate conceptualization can be expressed symbolically using a general linear model (McNeill, Newman & Fraas, 2012). For discussion purposes, consider adding an index, called an "affective index," as a criterion variable (not a predictor variable) to the calculation of value-added scores for teachers. In this multivariate approach Affective Index could represent a number of the non-school variables previously discussed. It could also include some school variables. For sake of simplicity, we add only one variable in this example.

Using LMEM (Field, 2005; Raudenbush & Bryk, 2002, Singer & Willett, 2003), the sequence of test scores for a student who is first tested in 1997 in the third grade is assumed to satisfy the following equations for testing between 1997 and 1999, from grade 3 to grade 5:

$$Y^{3}_{97} + AffectiveIndex^{3}_{97} = b^{3}_{97} + u^{3}_{97} + e^{3}_{97}$$
$$Y^{4}_{98} + AffectiveIndex^{4}_{98} = b^{4}_{98} + u^{3}_{97} + u^{4}_{98} + e^{4}_{98}$$
$$Y^{5}_{99} + AffectiveIndex^{5}_{99} = b^{5}_{99} + u^{3}_{97} + u^{4}_{98} + u^{5}_{99} + e^{5}_{99}$$

Where Y^{k}_{t} = test score in year *t*, grade *k*, $Affective\ Index^{k}_{t}$ = those variables determined relevant for determining a teacher's value-added score for year *i*, b^{k}_{t} = district mean test score in year *t*, grade *k*, u^{k}_{t} = contribution of the grade *k* teacher to the year u test score and e^{k}_{t} = student level components in year *t*, grade *k*. Building and classroom index are omitted here for simplicity (Ballou, Sanders, & Wright, 2004).

The teacher value-added scores for one year would then be calculated as follows:

$$(Y^{4}_{98} + AffectiveIndex^{4}_{98}) - (Y^{3}_{97} + AffectiveIndex^{3}_{97}) = (b^{4}_{98} + u^{4}_{98} + u^{3}_{97} + e^{4}_{98}) - (b^{3}_{97} + u^{3}_{97} + e^{3}_{97})$$

Since $u^{3}_{97} - u^{3}_{97} = 0$,

$$u^{4}_{98} = (Y^{4}_{98} - Y^{3}_{97}) - (AffectiveIndex^{4}_{98} - AffectiveIndex^{3}_{97}) - (b^{4}_{98} - b^{3}_{97}) - (e^{4}_{98} - e^{3}_{97})$$

The teacher effect (*u*) is what remains of the year-to-year gain after removing the district mean gain (*b*), the Affective Index gain, and the contribution of factors unique to the student (*e*). If we think of

$$(Y^{4}_{98} - Y^{3}_{97}) - (AffectiveIndex^{4}_{98} - AffectiveIndex^{3}_{97}) - (b^{4}_{98} - b^{3}_{97})$$

As the residual gain at the student level, quantifying teacher effects is a matter of determining how much of the residual gain to attribute to student specific factors, to the Affective Index, to the influence of the teacher or to school factors. Adding the Affective Index to the model will incorporate into the teacher effectiveness score a measure of student growth identified through

ways other than a single test score. It will more accurately reflect all the skills the classroom teacher addresses on a daily basis.

Options for Criterion Variables

Including the one factor, Affective Index, in the above multivariate approach to VAM is a simple example to demonstrate how a multivariate value-added model might be developed. The Affective Index could be a number of factors or constructs, both school and non-school. Many examples support the need to move to multidimensional VAM. First, Bandura's (1993) findings that students' perceived self-efficacy influences cognitive, motivational, affective, and selection processes could be a first step in determining the factors or constructs to include as criteria. Bandura states that "children's intellectual development cannot be isolated from the social relations within which it is imbedded or from its social consequences. It must be analyzed from a sociocultural perspective." (p. 137-138). In addition he explores the effect of teacher efficacy on student cognitive development, which may be quite important in student growth. Another example comes from Bar-On and Parker (2000). They document the importance of classroom experiences in contributing to student achievement. Further evidence comes from the conclusions of Hubbard and Datnow's (2005) ethnographic studies of single sex schools in California. Educators, they surmised, must "make emotional and social as well as academic knowledge explicit" (p. 128). Emotional and social learning goals are integral to school success. These and other research findings suggest student behaviors and teacher behaviors that impact cognitive development and that would produce the emergence of additional criterion variables that might be used in a multivariate approach to VAMs.

The most popular VAM, TVAAS, proposes that the school and non-school factors are accounted for in the VAM because the student is used as his/her own control. TVAAS does include the teacher effect in the model; in fact it is a persistent teacher effect over time. The reason more school/non-school effects are not included is that it is unlikely that the school/non-school factors for a child will change much over a year. It is assumed that the school/non-school effects are represented in the student level effects.

Franco (2006) analyzed VAS from Ohio (TVAAS model) to study their relationships among school/non-school factors. Correlation tables quantified relationships that exist between the independent variables (school/non-school factors) and dependent variables (value-added scores). For example, correlations for each independent variable, say percent of students with free and reduced lunch (%FRL), were determined for grade 3 reading value-added scores, for grade 3 mathematics value-added scores. Highly correlated factors were then used in a GLM to further study the relationships. Analyses revealed that some non-school and school factors do have significant relationships with the VAS. VAMs that do not include school/non-school factors may not be accurately representing student growth.

How might the vision of a multidimensional value added model be put into practice? When a school district is convinced that there is not one criterion variable, but many criterion variables, one scenario might suggest the following process. For example, the superintendent of the district could form a committee representative of all major stakeholders of the district. The committee might represent, for example, students, teachers, administrators, curriculum personnel, parents,

as well as stakeholders in the business, medical, legal, clerical, and social services sectors. Conducting focus groups with these stakeholders, a list of important criterion variables could be drawn up as the groups attempt to identify the important outcomes of their schools. Having accomplished that, they then generate the estimates (or measures) of each of those outcomes. As has been discussed earlier, criterion variables (educational outcomes) such as emotional intelligence, social skills, and self-efficacy, teacher-made exams, GPA, are possibilities. Once these variables are identified, student data on these measures form a dataset. The dataset could be factor analyzed to produce orthogonal constructs that represent the outcomes identified by focus groups. Because these dimensions are orthogonal, they are zero-correlated. The stakeholder focus group could weight each factor based on its importance to the community they represent. Perhaps factor 1, for example, is judged as twice as important as factor 2. Each factor is weighted relative to other factors. Next, a value added assessment is calculated on each factor separately. Each value added calculation is weighted according to the educational community's judgment. Combining these weighted scores would produce a composite value added score. To summarize what has been accomplished in this scenario: the school district has moved from a value added calculation using one criterion (standardized test scores) to a value added calculation using multiple criterion variables that have been identified and weighted by a group of stakeholders as the criteria they most highly value. The value added results are specific to this district because the criterion variables were locally identified and measured. Such a scenario increases the validity of the value added assessment tool to more thoroughly measure the educational outcomes of that district.

Summary

This proposal to consider a multivariate model for VAMs does not mean that VAMs are not functional, but that development of VAMs should be expanded to reflect what research has shown about teachers' impact on student progress. A test score alone does not accurately reflect student growth that could be a result of teacher/school or building interactions. In some districts, the VAS are incorporated into high stakes decisions such as teacher performance evaluation, salary increases or even restructuring of buildings. Educators from the classroom level to the district level criticize such high stakes decisions based on VAS because VAS fail to capture all facets of student growth. Educational researchers agree that VAS are not appropriate tools for high stakes decisions.

Academic intelligence is important for student success in further education and employment. Moreover, social and emotional intelligence, i.e., EQi (Bar-On, 2000), is also important for students to become contributing members of our society. Matthews, Zeidner, and Roberts (2002) explain that academic intelligence is essential to classroom success but everyday problem-solving requires practical intelligence as well. Measurement methodology requires that educators include other intelligences in value-added models that schools use to quantify school effects.

Specific building or district accountability ratings are used by the public to compare building or district effectiveness. As long as effectiveness ratings are based on a single test score, such comparisons may not be appropriate. For example, schools vary widely from district to district. Social contexts produce wide differences in students' lived experiences. Among other dynamics, schools vary widely in levels of parent involvement and financial support. The EQi's differ for

the student and community of learners in different types of districts. For accountability purposes, a VAS that incorporates more evidence about the students overall academic, social and emotional growth will be more helpful when determining teacher, building or district effectiveness than a VAS that incorporates only one test score.

Work should be done to determine what criterion variables should be added to the standardized test scores in calculating a more accurate and more valid value-added score. VAS that reflect the overall student growth and not just the academic test score growth will have more meaning to educators as well as to stakeholders.

References

Anderson, J. A., (2005). *Accountability in education*. Paris: International Institute for Educational Planning, & The International Academy of Education. Retrieved from http://www.unesco.org/iiep/PDF/Edpol1.pdf

Ballou, D., Sanders, W., Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics, 29*(1) 2004, 131-134.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117-148.

Bar-On, R., & Parker, J. D. A. (Eds.). (2000). *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school and in the workplace.* San Francisco: John Wiley & Sons.

Briggs, D. C., & Weeks, J. P. (2011). The persistence of school-level value-added. *Journal of Educational and Behavioral Statistics, 36*(5), 616-637.

Burnett, A., Cushing, E., & Bivona, L. (2012). Use of multiple measures for performance-based compensation.Washington, DC: Center for Educator Compensation Reform. Retrieved September 12, 2013, from http://0-cecr.ed.gov.opac.acc.msmc.edu/pdfs/CECR_Multi-pleMeasures.pdf

Caillier, J. (2010). Paying teachers according to student achievement: Questions regarding pay-for-performance models in public education. *Clearing House, 83*(2), 58-61.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood: Executive Summary*. Washington DC: National Bureau of Economic Research Working Paper No. 17699.

Field, A. (2005). *Discovering statistics using SPSS, 2nd ed.* Thousand Oaks, CA: Sage.

Franco, M. S. (2006). *The relationships among building level school/non-school factors and value-added scores in Ohio.* (Doctoral dissertation, University of Cincinnati, 2006). Embargoed by University of Cincinnati.

Glazerman, L., Loeb, S., Goldhaber, D., Staiger,D., Raudenbush, S., & Whitehurst, G. (2010). *Evaluating teachers: The important role of value-added*. Washington, DC: The Brown Center on Education at the Brookings Institute, 1-12.

Gong, B (2006, January 25). *Considering student growth for school accountability decisions: In the context of the NCLB growth models pilot.* Retrieved February 20, 2006 from http://www.ccsso.org/content/pdfs/StudentGrowthGong012506.ppt

Goldhaber, D. (2010*). When the Stakes Are High, Can We Rely on Value-Added?: Exploring the Use of Value-Added Models to Inform Teacher Workforce Decisions.* Center for American Progress. Retrieved on September 12, 2013 from http://cdn.americanprogress.org/wpcontent/uploads/issues/2010/12/pdf/vam.pdf

Good, R. H., & Kaminski, R. A. (2002). *DIBELS oral reading passages for first through third grades* (Technical Report No. 10). Eugene, OR: University of Oregon.

Hibpshman, T. (2004). *A review of value-added models.* Frankfort, KY: Kentucky EducationProfessional Standards Board.

Hill, H. C., Kapitula, L., & Umland, K. (2011). A validity argument approach to evaluating teacher value-added scores. *American Educational Research Journal*, *48*(3), 794-831.

Hubbard, L., & Datnow, A. (2005). Do single-sex schools improve the education of low-income and minority students? An investigation of California's public single-gender academies. *Anthropology and Education Quarterly, 36*(2), 115-131.

Jerald, C. (2009). *The value of value added data.* Educational Trust, 1-7.

Kane, T.J., Staiger, D.O. (2012). Gathering feedback for teaching: Combining high-quality observations with student survey and achievement gains. Seattle, WA: Bill & Melinda Gates Foundation. Retrieved June, 6 2013. From http://www.metproject.org/ downloads/MET_ Gathering_ Feedback_Practioner_Brief.pdf

Linn, R. L., & Miller, M. D. (2005). *Measurement and assessment in teaching, 9th ed*. Upper Saddle River, NJ: Pearson.

Matthews, G., Zeidner, M., & Roberts, R. D. (2002). *Emotional Intelligence: Science and Myth.* Boston: Massachusetts Institute of Technology.

McNeil, K. A., Newman, I., & Fraas, J. (2012). *Designing general linear models to test research hypotheses.* Lanham, MD: University Press of America.

Murphy, D. (2012). *Where is the Value in Value-Added Modeling?* [White paper]. Retrieved October 15, 2013 from, Pearson: http://educatoreffectiveness.pearsonassessments.com/downloads/viva_v1.pdf

Nitko, A. J. (2004). *Educational assessment of students, 4th ed*. Upper Saddle River, NJ: Pearson. No Child Left Behind Act, 20 U. S. C. § 6301(2001).

Papay, J. P. (2011). Different tests, different answers: The stability of teacher value-added estimates across outcome measures, *American Educational Research Journal, 48*(1), 163-193.

Perez-Pena, R. (May 29, 2012). Waivers for 8 more states from 'No Child Left Behind.' Retrieved: http://www.nytimes.com/2012/05/30/education/eight-more-states-get-waiver from-no-child-law.html?_r=1

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods, 2nd ed.* Thousand Oaks, CA: Sage.

Sanders, W. L., & Horn, S. P. (1994). The Tennessee value-added assessment system (TVAAS) mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education, 8,* 299-311.

Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee value-addeda ssessment system (TVAAS) Database. Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12*(3), 1998. 247-256.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford: Oxford University Press.

Timmermans, A C., Doolaard, S., & de Wolf (2011). Conceptual and empirical differences among various value-added models for accountability. *School Effectiveness and School Improvement, 22*(4), 393-413.

van de Grift, W. (2009). Reliability and validity in measuring the value added of schools. *School effectiveness and School Improvement, 20*(2), 269-285.

Yeh, S. T. (2012). The reliability, impact, and cost-effectiveness of value-added teacher assessment methods. *Journal of Education Finance, 37*(4), 374-399.

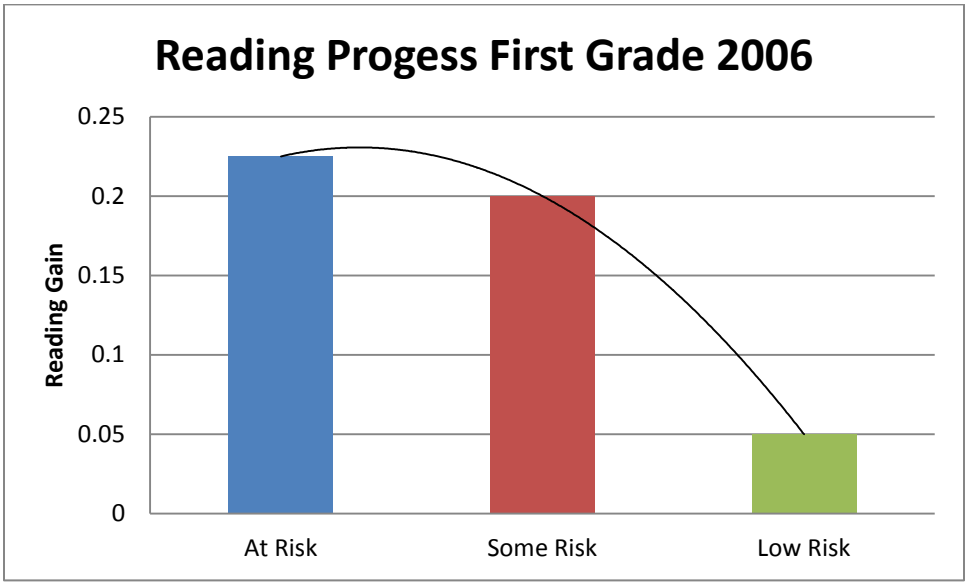


Figure 1. Typical Shed Pattern demonstrating most gains in the "at risk" group in the Reading First Ohio Center (RFOC) study.

.

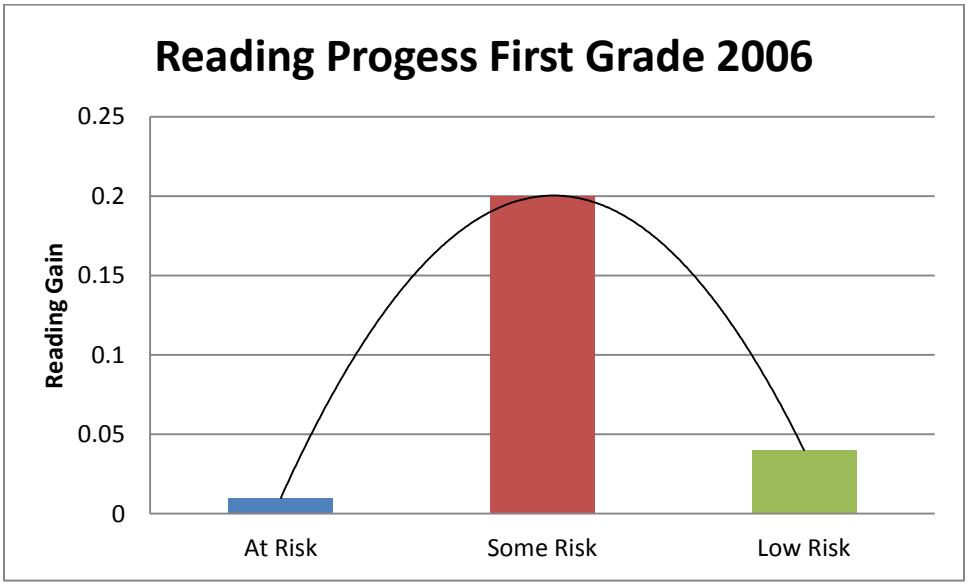


Figure 2. Typical Tepee Pattern demonstrating most gains in the "some risk" group in the Reading First Ohio Center (RFOC) study.

.

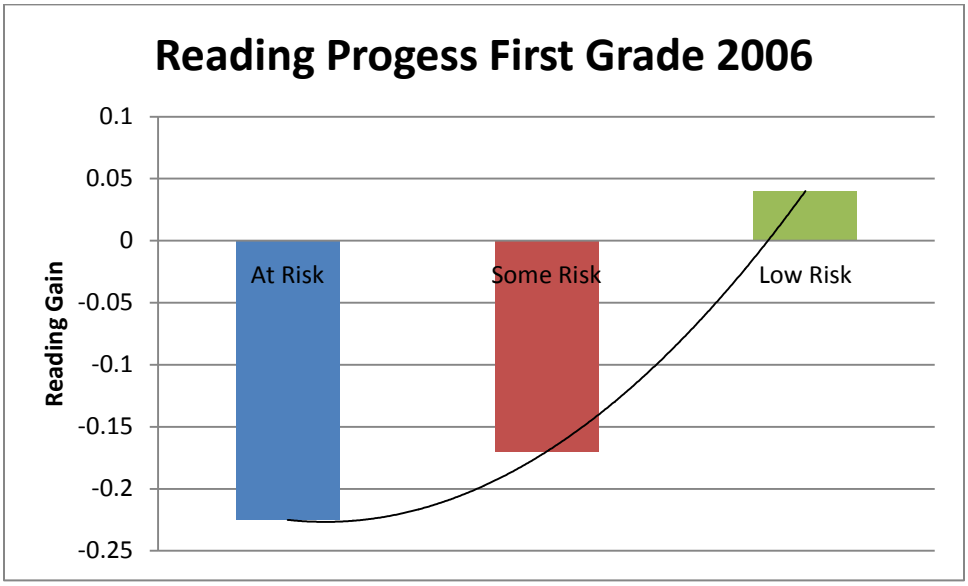


Figure 3. Typical Upward Shed Pattern demonstrating most gains in the "low risk" group in the Reading First Ohio Center (RFOC) study.

Gender Differences in Resilience of Academic Deans

Albert J. Isaacs
University of Phoenix

Abstract

The purpose of this investigation was to determine the difference in the levels of resilience characteristics between male and female deans within a state university system. Resilience is the ability to operate in a changing environment while consistently maintaining one's effectiveness. This quantitative study utilized the survey, Personal Resilience Questionnaire(PRQ). Responses were received from 35 deans, 16 males and 19 females, and the results showed that the average means scores suggested that female deans had a higher level of resilience. The Mann-Whitney U-test showed that only one of the resilience characteristics – Proactive (.046) – was significant at the .05 level. Numerous studies have shown that females tend to face more difficulties within the workplace (e.g. Trentham & Larwood, 2004; Reskin, 1993). Presumably the accumulation of these challenges has enhanced resilience for women. The identification of resilience levels within academic administration can assist university leaders to operating effectively during times of change.

Leadership in higher education is a topic that has been widely studied by scholars. The leadership characteristics of university presidents, provosts, academic deans, department chairs, and faculty have all been studied, illustrating their role within the institution and how their leadership impacts the educational setting. One aspect of higher education leadership that has not been adequately researched, however, is the concept of gender differences in resilience. The ability of individuals to be resilient is important in the field of higher education as the constantly changing environment provides many challenges and adversities for administrators to overcome.

Higher education today is operating in a constantly changing environment. Advances in technology, mission creep, changes in educational policy, financial limitations, changes in student demographics, and the increased call for accountability are just some of the factors that have required the field of higher education to respond in a manner it has never had to before (Del Favero, 2005; Lucas, 2000; Gmelch & Miskin, 1993), and its constituents to respond accordingly. This constant change creates an environment where it can be difficult for higher education professionals to be successful. To overcome these challenges, it then becomes important for academic administrators to be effective in operating within such an environment and have the skills necessary to be successful despite adversities.

Astin and Astin (2000) spoke to the turbulent and changing environment that encompasses today's society, and to higher education's role and responsibility to help facilitate and manage that change. Though they spoke more of societal problems such as global warming, religious and ethnic conflict, the increasing ineffectiveness of government, and the misdistribution of wealth and opportunity, they underlined the fact that transformative leadership stemming from the university and its teachings is a key component to bringing about positive change. Astin and

Astin pointed out that higher education plays a critical role in training and shaping leadership in society, and further situates higher education's presence in a constantly-changing environment. Operating within an external environment that experiences persistent and continuous change, it is inevitable that higher education will have to contend with change as well. This is not only a by-product of its response to the external environment, but also of its own change management issues. The influence that change has on the field of postsecondary education makes it a factor that higher education professionals need to be cognizant of, with an intentional, stated objective of being successful and productive despite the challenges posed by a changing environment.

Trowler (1998) gave specific examples of how academe had responded to "rapid changes in higher education." Though he concentrated primarily on the credit framework – in itself a response to changing needs in higher education – Trowler also alluded to the size of the system, the composition of the student body, available resources, the purpose of higher education, and future objectives as being factors that had been altered in recent years to respond to change. Individually, each of these factors is a critical component to any system of higher education. Taken as a unit, however, changes in each factor could result in a shift in the postsecondary education paradigm.

Though changes have been made to respond to these factors, it is difficult for higher education to react quickly enough. Walker and Salt (2006) declared that humans are typically good at acknowledging and responding to rapid change, but not as good at responding to things that may change slowly. It can be argued that many of the changes that affect higher education today are slow, gradual changes, such as increases in enrollment, changes in student demographics, and institutional culture.

Traditionally, higher education has been slow to respond to change, but the increasing acknowledgment and acceptance that change is a prevailing force which must be addressed has resulted in administrators being charged with an increased responsibility to perform effectively within change. Kittelson and Transue (1984) outlined four centuries of history in the creation and development of universities, and illustrated numerous external factors which threatened the continued existence and success of postsecondary institutions. Generational differences were the major form of adversity, but universities and its members have managed to overcome these challenges despite the constant changing environment. Spitz (1984) explained this by stating that "universities are tough, resilient institutions capable of surviving dormant periods, hostile forces, and even then of emerging as revitalized centers of new learning" (1984, p. 63). This notion of universities being resilient and being able to transcend adverse situations would imply that ideally university staff be resilient as well.

Due to current demands, professionals must be successful in leading their organizations through change by having the knowledge and skills necessary to do so. It is important for university administrators to be well-equipped and effective at implementing and maintaining this positive change, and for them to exhibit these characteristics in such a manner that other members of the organization reflect similar behavior.

While external forces can create many challenges to working in higher education, university administrators are faced with many internal adversities as well. Jacobs, Cintron, and Canton

(2002) spoke to the challenging environment associated with working in academe, providing narratives of gender diverse scholars who have persevered to be successful in American academia. The authors presented examples of some traditional challenges associated with succeeding in academia, such as evaluation, tenure, an understanding of internal politics, and scholarly expectations of rigor, quality, and productivity. The primary focus of this work is the achievement of faculty in "retaining their self-identity and self-respect in the face of prejudice, biases, and disrespect from colleagues, peers, and administrators" (p. 9) but the authors also illustrated that resilience is an important attribute for any scholar looking to 'survive' in academe, particularly those from underrepresented groups.

The issue of change in higher education is not a new one, however, as these internal and external factors have created a challenging working environment for higher education professionals for many years. Faculty, staff, and students have had to adjust to evolving cultures and environments since the initial inception of colleges and universities, and scholars have documented this change. The era that we operate within today, however, is creating a new culture that affects higher education on numerous levels, posing new challenges such as budgetary constraints, the rapid expansion of information, virtual universities, technology, emphasis on diversity, and an increased call for accountability from higher education institutions and its constituents (Lucas, 2000). Gmelch and Miskin (1993) also identified changing student demographics, disintegrating college curricula, and shifting attitudes and practices of faculty as being major challenges that higher education has to cope with. The culmination of these challenges creates an environment that can be difficult to operate effectively within.

To counteract the challenges posed by a changing environment, Hiatt and Creasey (2003) and Luecke (2003) discussed numerous change management principles that organizations can employ to operate efficiently during times of change, including organizational assessments, sponsor preparation, value systems, and the ADKAR (Awareness, Desire, Knowledge, Ability, and Reinforcement) model. Among these are organizational and individual resilience – with resilience defined as "the capacity to absorb high levels of change while displaying minimal dysfunctional behavior" (Conner, 1992, p. 219) – which are key constructs that can assist institutions of higher education to be successful in a changing environment.

Organizational leaders' ability to be resilient during change and professional challenges, and how they can use that resilience to effectively guide the organization and its members towards institutional goals and objectives, is critical given this environment of constant change (Brooks & Goldstein, 2003; Reivich & Shatte, 2002; Deevy, 1995; Conner, 1992). Resilience is a tool that higher education professionals can use to be successful in a changing environment, and to assist the field in its response to both internal and external factors. The proper development and utilization of the concept of resilience can provide benefit to organizations, is the cornerstone of this study. The ability to operate in a changing environment while consistently maintaining one's effectiveness as a higher education professional, can have meaningful results to the success of the organization as a whole. Having resilient academic administrators in place to lead the organization during difficult times can help alleviate some of the challenges associated with such a dynamic environment.

The purpose of this investigation was to determine the difference in the levels of resilience characteristics between male and female academic deans within a state university system. This quantitative study utilized the survey method in its research design involving an instrument that evaluates and measures resiliency, which was sent to the target population of academic deans via e-mail.

The guiding research question in this study is: What are the differences in the levels of resilience characteristics between male and female academic deans? The following were the hypotheses for this quantitative study: ***H1o (Null Hypothesis):*** There are no significant differences in the levels of resilience characteristics between male and female academic deans. ***H1a (Hypothesis):*** There are significant differences in the levels of resilience characteristics between male and female academic deans.

Method

Participants

The participants for this study was all academic deans employed at public universities in a large state university system. At these institutions, there were a total of 87 academic deans, who provided the main source of data for this study.

Sampling Procedure

The presidents of each institution were contacted to garner support for the study and to obtain permission to contact their academic deans. Of these institutions, all but two granted permission to proceed with the study. For those institutions that granted permission for their staff members to be contacted, all 87 deans recognized as such by their respective institutions were surveyed, regardless of the academic area they worked in, including those of professional schools, undergraduate studies, and graduate colleges.

Respondents were asked to complete the survey by a specific deadline. Once this deadline passed, the survey request was sent out two additional times in an effort to garner more responses and achieve statistical validity.

Data Collection Instrument

The data collection instrument used in this study is a survey developed by Daryl Conner and his company, formerly known as ODR-USA, Inc. The Personal Resilience Questionnaire (PRQ) measures individuals' resilience across each of the seven key resilience characteristics: being positive about the world, positive about themselves, focused, flexible in their thoughts, flexible towards others in their social environment, organized, and proactive, and is comprised of 75 questions. All questions are based on a six-point Likert-type scale, measuring the individual's response to each situation as being: strongly disagree, disagree, slightly disagree, slightly agree, agree, and strongly agree.

Respondents' answers to these questions result in scores for each of the seven resilience characteristics, with each score based on a 100-point scale. This score is a percentile measuring

their results against all persons who have taken the Personal Resilience Questionnaire (PRQ), which is a population of over 64,000 people and include individuals from a broad composition.

Conner Partners administered the survey and agreed to tailor the first part of the *Personal Resilience Questionnaire* to include the demographic question of the number of male and female academic deans.

Conner Partners were the only organization capable of scoring the *Personal Resilience Questionnaire* and the responses and they agreed to facilitate the collection of the data. Once all submissions had been received, Conner Partners provided the researcher with the results for the analysis.

Data analysis

Means were calculated for each resilience characteristic and comparisons made to evaluate which gender position exhibited higher results. Additionally, the Mann-Whitney *U* test was used to test whether differences between the two populations (female and male deans) were significant. The Mann-Whitney *U* test is a non-parametric test for testing and estimating differences between two populations and is useful for populations with arbitrary sample sizes and shapes (Hettmansperger, 1991).

Results

This study measured the resilience of male and female academic deans at institutions of higher education. At these institutions, there were a total of 87 deans. Responses were received from 35 deans, for a response rate of 40.2%. Of the 35 academic deans who participated in the study, 16 were male and 19 were female. The mean scores for deans were as follows: See Table1.

Table 1
Mean Scores for Female and Male Deans for Each Resilience Characteristic

	Female	Male
Positive: World	74.37	70. 69
Positive: Self	83.79	80.63
Focused:	77.84	67.44
Flexible: Thoughts	80.26	65.44
Flexible: Social	74.58	62.75
Organized	57.84	51.75
Proactive	79.05	61.19

The average scores shown in Table 1 illustrate that female deans in this study had a higher level of resilience in each of the seven resilience characteristics than did male deans, indicating that females, on the average, tend to be more resilient.

Although females averaged a higher level of resilience in each of the seven characteristics, the Mann-Whitney *U* test employed indicated a significant difference on one of the seven resilience characteristics between male and female deans.

Results of the Mann-Whitney *U* test are found in Table 2 below.

Table 2
Mann-Whitney U Test for Differences Between Female and Male Deans

	Mann–Whitney U- test	Z-scores	Asymp. Sig(2 tailed)
Positive: The World	140.500	-.383	.702
Positive: Self	133.000	-.636	.525
Focused	110.000	-1.404	.160
Flexible: Thoughts	107.000	-1.495	.135
Flexible: Social	109.500	-1.413	.158
Organized	131.000	.697	.486
Proactive	92.000	-1.994	.046

The Mann-Whitney test in Table 2 shows that although female deans in this study averaged a higher level of resilience on all seven resilience characteristics than did male deans, only one of the resilience characteristics – *Proactive* (.046) – was significant at the .05 level.

In conclusion, based on the results of mean scores, the researcher answered the research question that there are differences in levels of resilience characteristics between male deans and female deans. Additionally, based on the results of the Mann-Whitney U test, the researcher accepted the hypothesis, that there are a significant difference in the levels of resilience characteristics between male and female deans.

Discussion

It is noteworthy that female academic deans exhibited higher levels of resilience than male academic deans in each of the seven characteristics *(Positive: The World, Positive: Yourself, Flexible: Thoughts, Flexible: Social, Focused, Organized and Proactive)* with one of those characteristics *(Proactive)* being significant, as that supports the hypothesis that there is a significant difference in the levels of one resilience characteristic, *Proactive*, between male and female academic deans.

This hypothesis was based on the notion that higher levels of resilience prepares individuals for higher levels of success. Numerous studies have shown that females tend to face more difficulties within the workplace (e.g. Trentham & Larwood, 2004; Reskin, 1993). Presumably

the accumulation of these challenges has enhanced resilience for women, or it may just have been necessary for women to be more resilient to compete for the position at each level. Researchers have recognized the changing environment of higher education and the impact that can have on the success of deans, and are looking for new competencies that can guide these administrators through these difficulties.

The significant characteristic of resilience, *Proactive,* indicates that resilient individuals tend to challenge the world around them, tending toward active rather than reactive approaches to problem solving" (ODR, n.d.). Failure is viewed as an opportunity to learn, and in the face of adversity strive to take active strategies rather than use avoidance and withdrawal strategies. This decision to view risks as desirable enhance adaptation effectiveness by leading people to set high standards, which can lead to high performance, and their assertiveness may detect early signs of "potential changes and discrepancies and the facilitation of quick and effective responses (ODR, 1995).

Resilience is a key construct of leadership in a changing world and with so many individuals looking at you to lead, resilience is important to be effective. Having an understanding of one's resilience is beneficial not just because it shows the areas where you may be more effective, but more so because it identifies areas that can be improved upon. Enhancing resilience could be a key thing to elevate one's performance above these challenges and provide the skills needed to be effective despite these obstacles.

The identification of resilience levels within academic administration can assist university leaders in understanding whether their institutions are maximizing their potential for operating effectively during times of change.

References

Brooks, R., & Goldstein, S. (2003). *The power of resilience.* New York: McGraw-Hill.

Conner, D. R. (1992). *Managing at the speed of change.* New York: Villard Books.

Deevy, E. (1995). *Creating the resilient organization: A rapid response management program*. Englewood Cliffs, NJ: Prentice Hall, Inc.

Del Favero, M. (2005). The social dimension of academic discipline as a discriminator of academic deans' administrative behaviors. *The Review of Higher Education, 29*(1), 69-96.

Gmelch, W. H., & Miskin, V. D. (1993). *Leadership skills for department chairs*. Bolton, MA: Anker Publishing Company.

Hettmansperger, T. P. (1991). *Statistical inference based on ranks*. Malabar, FL: Krieger Publishing Company.

Jacobs, L., Cintron, J., & Canton, C. E. (Eds.). (2002). *The politics of survival in academia: Narratives of inequity, resilience, and success.* Lanham, MD: Rowman & Littlefield Publishers.

Kittelson, J. M., & Transue, P. J. (Eds.). (1984). *Rebirth, reform and resilience: Universities in transition 1300-1700.* Columbus, OH: Ohio State University Press.

Lucas, A. F. (2000). *Leading academic change: Essential roles for department chairs*. San Francisco: Jossey-Bass.

ODR. (1996). *Criterion-related validity of the personal resilience questionnaire*. Unpublished Report by ODR, Inc. at Atlanta, GA.

ODR. (1995). *Change resilience: A cognitive research approach*. Unpublished Report by ODR, Inc. at Atlanta, GA.

Reivich, K., & Shatte, A. (2002). *The resilience factor: 7 essential skills for overcoming life's inevitable obstacles*. New York: Broadway Books.

Resnick, M. (2000). Protective factors, resiliency, and healthy youth development. , Philadelphia: Hanley & Belfus, Inc.

Rutter, M. (1985). Resilience in the face of adversity. Protective factors and Resistance to psychiatric disorder. *British Journal of Psychiatry, 147*, 598-611.

Spitz, L. W. (1984). The importance of the reformation for the universities: Culture and confessions in the critical years. In J. M. Kittelson and P. J. Transue (eds.), *Rebirth, reform and resilience: Universities in transition 1300-1700* (pp. 42-67). Columbus, OH: Ohio State University Press.

Trowler, P. R. (1998). *Academics responding to change: New higher education frameworks and academic cultures.* Philadelphia, PA: SRHE Open University Press.

Trentham, S., & Larwood, L. (2004). Gender discrimination and the workplace: An examination of racial bias theory. *Sex Roles, 38*(1-2), 1-28.

Walker, B., & Salt, D. (2006). *Resilience thinking: Sustaining ecosystems and people in a changing world.* Washington, DC: Island Press.

What is Cheating? Student and Faculty Perception of what they Believe is Academically Dishonest Behavior

George Watson
Marshall University

James Sottile
Missouri State University

Jia Grace Liang
University of Georgia

Correspondence related to this article should be directed to Dr. George Watson at Watson@marshall.edu

Abstract

The study of ethics and moral development of college students is an important issue. Knowing and understanding the ethical behavior of college students can lead to changing and increasing appropriate behavior among graduate and undergraduate students. Such changes in ethical behavior and moral development during the college experience can strengthen the foundation for appropriate adjustments and foster a greater awareness for positive ethical behavior throughout a lifetime. This research study examined the perception of what students and faculty believe is academically dishonest behavior by identifying different types of scenarios. Given the cheating behavior by students, it is important to know what students and faculty actually believe is academically dishonest behavior. The research question was "What do students and faculty perceive as cheating?" Students and faculty were surveyed and the findings indicate a clear discord between perceptions of cheating and actual cheating as determined by students and faculty.

The issue of cheating in academic situations has been studied regarding what students believe but faculty beliefs in this area are more limited. Graham (1994) reviewed both faculty and student beliefs about cheating. 480 students and 48 faculty completed a survey and 89% admitted to cheating (Graham, 1994). Graham (1994) noted that "attitudinal variables were better at predicting cheating than were background variables" (p. 255). Roig and Ballew (1992) also completed a study that reviewed faculty and student attitudes about cheating. It was found that student perceptions of faculty beliefs about cheating were similar to what faculty actually believed but the same was not true regarding the perceptions that faculty had about student cheating. Faculty believed that students were more liberal in their understanding of cheating but the students did not have this same belief about their cheating behavior. Business related majors were the most tolerant of cheating behaviors. When 364 engineering students and 80 faculty were surveyed, 62% of students admitted to copying homework but only 51% of faculty thought this was cheating and 56% of students admitted to cheating (Singhal, 1982). When faculty syllabi were reviewed, Volpe, Davidson, and Bell (2008) found no relationship between the number of integrity related statements in the syllabus and attitudes about student cheating

behaviors. The amount of cheating was underestimated by faculty and the amount of cheating that occurs does not correspond with written guidelines (Volpe, Davidson, & Bell**,** 2008).

Cheating, or academic dishonesty, has been found to be common in studies over several decades and has raised concerns on college and university campuses more than ever before (Bowers, 1964; Covey, 2008; Gulli, Kohler, & Patriquin, 2007; Kleiner & Lord, 1999; McCabe, 1992 and 1997; McCabe, Treviño, & Butterfield, 2001; Rimer, 2003). In a nationwide research study of 23 public and private colleges and universities, McCabe (2001-02) found that 38% of the undergraduate student surveyed indicated that they had engaged in Internet plagiarism (cited in Rimer, 2003). A survey of 5300 U.S. graduate students at the Academy of Management Learning and Education found business students, at 56%, were the worst offenders, followed by engineering students, at 54%, in the engagement of unethical behaviors from plagiarism to using unauthorized notes in exams (Gulli et al., 2007). According to *The Chronicle for Higher Education*, in November 2010, more than 200 of the 600 students in a University of Central Florida business class admitted that they benefited from accessing online test questions prior to taking their midterm exam (The Ticker, 2010). Brown, Weible, and Olmosk (2010) observed that 100% of the students in an undergraduate management class in 2008 admitted to cheating versus 49% of students in undergraduate marketing classes. Academic cheating has also involved alumni. "Two students and an alumnus from Florida International University were arrested on felony charges for stealing a test by hacking into a professor's computer, reports the Sun Sentinel" (Wiley Periodicals, 2014, p. 2).

Influences upon Cheating Behavior

Many factors can influence cheating behavior. An individual's traits and characteristics can affect his or her morality (Kanfer, Wanberg, & Kantrowitz, 2001; McCabe, 1997; Shipley, 2009). Machiavellianism, for example, is "an individual difference characteristic that focuses on the extent to which individual hold cynical views of human nature, behave manipulatively in their interactions with others, and generally have a low regard for traditional or conventional standards of morality" (Christie & Geis, 1970, cited in Bloodgood, Turnley, & Mudrack, 2010, p. 26) and it has been found to be negatively related to ethical awareness and behavior (Bloodgood et al 2010; Bolino & Turnley, 2003; Granitz, 2003; O'Fallon & Butterfield, 2005; Tang & Chen, 2008). Studies of ethical conduct also found gender related differences. Females, in general, demonstrate higher ethical standards than males (e.g., Borkowski & Ugras, 1992; Humbarger & DeVaney, 2005; Shepard & Hartenian, 1991). Stevenson (1999), for example, found that females reported significantly higher cognitive moral judgment scores than males. Nevertheless, Lester and Diekoff (2002) noted that the majority of traditional cheaters are women whereas a majority of on-line cheaters are men. Age also plays a role in a student's ethic decision-making process. A student's ethical values increase with his or her age (Humbarger & DeVaney, 2005; Ruegger & King, 1992). Contradictorily, researchers studying babies and young toddlers at the Yale Infant Cognition Center and other institutions such as Harvard suggest that morality is a trait endowed with us at birth, and this "infant morality" turns more selective as we grow – in other words, we are losing some positive social inclination as we are socialized by the culture(s) we live in (Tucker, 2013). Whether the student is extrinsically or intrinsically motivated also plays a role in students' engagement in academic dishonesty. Rettinger and Kramer (2009) found that students engaging in unethical behavior were extrinsically motivated.

Different from the individual differences approach, the other camp of scholars focuses on the contextual factors that influence students' decisions to cheat (or correlates of cheating). Scholars of the *Theory of Planned Behavior* (TPB) suggested that students' cheating behavior is influenced by (1) attitude toward cheating, (2) perceived social pressures to engage or not engage in cheating; and (3) the perceived ease of performing cheating (Ajzen, 1991, 2002; Genereux & McLeod, 1995; Nonis & Swift, 2001; Passow, Mayhew, Finelli, Harding, & Carpenter, 2006; Whitley, 1998). Graham (1994) noted that compared with other background variables, a student's attitude toward cheating is better at explaining his or her cheating behavior. Students with favorable attitudes toward cheating are more likely to cheat than those who have unfavorable attitudes (Nonis & Swift, 2001; Whitley, 1998). Neutralizing attitudes – "beliefs that an individual holds to justify cheating behavior" (Hsiao & Yang, 2011, p. 304) is essential to understanding cheating because any blame or guilt resulting from conducts of cheating can be counteracted or neutralized (Covey, 2008; Diekhoff et al., 1996; McCabe, 1992). Neutralized attitudes toward cheating cultivate a culture of cheating and explains why knowing it is wrong to cheat does not necessarily stop students from engaging in cheating behaviors (Baird, 1980; Davis, Grover, Becker, & McGregor, 1992; Haines, Diekhoff, LaBeff, & Clark, 1986; Pulvers & Diekhoff, 1999; Rettinger & Kramer, 2009). For example, students in a study conducted by Haines and her colleagues (1986) believed that cheating is a personal behavior and will not hurt anyone and thus it is acceptable. Furthermore, students' attitudes toward cheating vary along a number of dimensions, namely, the assessment type (e.g., exams or papers or homework), the intention (whether the misconduct is planned in advance or spontaneous), and the role (whether a student is providing or receiving assistance) (Grijalva, Nowell, & Kerkvliet, 2006; Hard, Conway, & Moran, 2006; Murdock, Beauchamp, & Hinton, 2008; Passow et al., 2006; Vitell & Muncy, 1992).

These distinctions are important to the extent that students view certain misconduct as cheating but not the others and consider certain misconduct as more serious than others (Bisping, Patron, & Roskelley, 2008; Bloodgood et al., 2010; Jones, 2011; Jordan, 2001; Lim & See, 2001). For example, students considered examination cheating more serious than plagiarism (Lim & See, 2001) but did not perceive turning in an assignment previously submitted for another class as plagiarism or cheating (Jones, 2011). Some researchers suggested that lack of knowledge about what constitutes academic dishonesty contributes to this confusion (Blum, 2009; Carroll, 2007; Hansen, 2003; Howard & Davies, 2009). There are studies finding many students' academic dishonesty related to Internet use as the result of their belief that Internet information is public and free from intellectual property rights and thus failing to cite internet sources is not cheating (Ma, Wan, & Lu, 2008; McCabe, 2001-02, cited in Rimer, 2003; Schrimsher, Northrup, & Alverson, 2009).

Other contextual factors are found to be influential to students' cheating behaviors. Studies indicate that the level of cheating differs by college majors (Baird, 1980; Bowers, 1964; Jackson, Levine, Furnham, & Burr, 2006; McCabe, 1997; Newstead, Franklyn-Stokes, & Armstead, 1996; Rawwas & Isakson, 2000; Shaughnessy, 1988) and the highest percentage of undergraduates reporting cheating are those enrolled in "vocationally oriented majors such as business and engineering" (McCabe, 1997, p. 444). The differences have implications for the effectiveness of ethics education in various academic disciplines (King & Mayhew, 2002; Luthar & Karri, 2005;

Williams & Dewett, 2005). Peers were also found to be influential in students' attitudes toward cheating. Observation and/or perceptions of others' cheating encourage students to cheat as well (Bowers, 1964; Gulli et al., 2007; Hard et al., 2006; Koljatic & Silva, 2002; Teodorescu & Andrei, 2009; Watson & Sottile, 2010; Whitley, 1998). Students' perceptions of the quality and relevancy of instruction also influence their cheating behaviors (Okoro, 2011; Teodorescu & Andrei, 2009). When satisfaction with faculty's instruction declines, it creates "desperation and tension" (Okoro, 2011, p. 177) and "students may well devalue it, making it easier to justifying cheating" (Teodorescu & Andrei, 2009, p. 281).

Given the limited amount of research that examines both student and faculty observations of academic cheating, this research study examined the perceptions of what students and faculty believe is academically dishonest behavior by identifying different types of scenarios. Given the cheating behavior by students, it is important to know what students and faculty actually believe is academically dishonest behavior. The research question was "What do students and faculty perceive as cheating?"

Method

Sample and Participant Selection

The survey was administered to 400 undergraduate/graduate students and 57 faculty. The student gender breakdown was 122 males, 276 females, and 2 that did not identify a gender, while the faculty division was 32 male and 25 female.

Assessments and Measures

The instrument was divided into three parts, a demographic section of three questions, and a section in which participants were given 20 scenarios and asked to identify whether they believed the scenario represented academically dishonest behavior. A third section, in which respondents were given the same 20 scenarios and asked if they should be in a new academic dishonesty study was not used in this research. "Yes" responses were given a value of 1 and "No" responses were given a value of 2. These scenarios were single sentence statements covering a wide range of possibilities, from using study guides to seducing classmates for help. The Cronbach Alpha reliability measure for the instrument was .824. The surveys were delivered to participating students in classrooms by one of the researchers, collected after participant completion, and held by investigators. The faculty survey was identical but with two fewer demographic questions dealing with academic rank and college but was otherwise the same.

Sampling Procedures

Participation requests to campus instructors at a Mid-Atlantic university were emailed by the investigators to ask permission to come to class and give the survey instrument, and the student participants were selected from classes in which the instructor had volunteered to let students take the survey. Students in these classes could choose to opt out and not complete the instrument without repercussion. Faculty were invited to participate in the research via email

during the opening of the fall semester of 2013. Faculty followed a link to the survey, completing the same instrument minus two demographic questions, "College rank (Freshman, Sophomore, Junior, Senior, or Graduate)", and "Major College."

Results

Exploratory Factor Analysis

The factor analysis of the responses yielded three factors that accounted for 62% of the variability of responses, These factors were students who had someone else complete or help with the assignment, technical issues/resubmitting previous own work, and studying from previously created materials. Having someone else complete the assignment was defined as situations in which work completed by someone other than the student was submitted as the student's own. This included such scenarios as having a friend complete homework, paying for a term paper, or receiving test answers while taking an exam. Technical issues were, for example, lying to the professor to get a time extension or submitting your own work from other courses instead of creating something original. Studying from previously created materials encompassed scenarios such as studying course exams from previous semesters or using others' notes. The Cronbach Alpha coefficients for each factor ranged from .638 to .756. Table 1 shows the factors with related scenarios and Table 2 yields the means and standard deviations of the factors.

Analysis of Variance (ANOVA)

The analysis of variance (ANOVA) results for student/faculty data yielded significant results ($p < .05$) for several scenarios as outlined in Table 3. In all but two scenarios faculty mean score for cheating perception was lower, indicating that faculty felt the behavior was cheating more so than students. Two exceptions to this were the statements, "You review exams taken by friends in previous semesters to study." ($F_{1, 446} = 3.918$, $p = .048$) and, "You study using the study guide provided by the professor." ($F_{1, 454} = 8.516$, $p = .004$). Faculty had a higher mean score than students in both cases. Table 3 gives the analysis of variance between students and faculty responses to all survey questions.

Discussion

The results of this study show that there are significant differences between students and faculty in what constitutes academic dishonesty. This difference puts the issue of cheating as more than a simple crime and punishment but also as a teaching issue. If the argument is made that dishonest behavior is on the rise, data from this study would indicate that students are not becoming more immoral but simply more uneducated in what is acceptable. Part of this issue is directly related to the advent of new technologies and social media. Students today have grown up in a world in which access to information is literally at their fingertips, communication with friends is instantaneous, and the separation of class time and social time is blurred. Less than a generation ago students were physically separated from the outside world within the walls of the classroom and dishonest behaviors were limited to cheat sheets, plagiarism, and paying others to write term papers. Today, cell phones, computers, and tablet devices remove the physical

barriers of the classroom and create opportunities for students to receive information and assistance instantly.

At the same time we question the ethics of students and worry about the different ways students can receive assistance within our educational system. This system promotes critical thinking and team work but the issue of ethical and moral education needs to be part of this process. Educators need to avoid sending confusing or mixed message related to what is acceptable behavior in both K-12 and higher education classrooms. In addition, some self-examination of how we assess learning would be beneficial to insure our teaching and learning systems are still the best models in today's educational world.

Character education is a concept frequently heard within the K-12 arena. Observances from this study indicate that this concept needs to be expanded to formal education beyond high school years. Universities and colleges need to have at least one class per major related to the process of moral/ethical resolution. Cheating is an unethical behavior reported by many students. In a technological world where access to information is at the click of the mouse button, "cheating" in many formats is effortless. Professors must not only stress the content requirements of an assignment, but also the ethical responsibilities of doing such assignments. Professors must recognize that any assignments requiring out-of-class work will quite probably reflect group work rather than individual effort, and if non-collaborating students are aware of this activity, they are unlikely to report it to the professor. This means that in courses that stress the measurement of individual achievement, more classroom time will have to be dedicated to this activity. And even during those classroom activities, the professor will have to remain vigilant.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*, 179-211.

Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology, 32*(4), 665-683.

Baird, J. (1980). Current trends in college cheating. *Psychology in the Schools, 17*(4), 515-522.

Bisping, T. O, Patron, H., & Roskelley, K. (2008). Modeling academic dishonesty: The role of student perceptions and misconduct type. *Journal of Economic Education, 39*(1), 4-21.

Bloodgood, J. M., Turnley, W. H., & Mudrack, P. E. (2010). Ethics instruction and the perceived acceptability of cheating. *Journal of Business Ethics, 95*(1), 23-37.

Blum, S. D. (2009). Academic integrity and student plagiarism: A question of education, not ethics. *The Chronicle of Higher Education, 55*(24), A35.

Bolino, M. C., & Turnley, W. H. (2003). More than one way to make an impression: Exploring profiles of impression management. *Journal of Management, 29*(2), 141-160.

Borkowski, S., & Ugras, Y. (1992). The ethical attitudes of students as a function of age, sex, and experience. *Journal of Business Ethics, 11*(12), 961-979.

Bowers, W. J. (1964). *Student dishonesty and its control in college*. New York, NY: Bureau of Applied Social Research, Columbia University.

Brown, B. S., Weible, R. J., & Olmosk, K. E. (2010). Business school deans on student academic dishonesty: A survey. *College Student Journal, 44*(2), 299-308.

Carroll, J. (2007). *A handbook for deterring plagiarism in higher education* (2nd ed.). Oxford, England: Oxford Center for Staff and Learning Development, Oxford Brookes University.

Christie, R., & Geis, F. (1970). *Studies in Machiavellianism*. New York, NY: Academic Press.

Covey, S. (2008). *The speed of trust*. New York, NY: Free Press.

Davis, S. F., Grover, C. A., Becker, A. H., & McGregor, L. N. (1992). Academic dishonesty: Prevalence, determinants, techniques, and punishments. *Teaching of Psychology, 19*(1), 16-20.

Diekhoff, G. M., LaBeff, E. E., Clark, R. E., Williams, L. E., Francis, B., & Haines, V. J. (1996). College cheating: Ten years later. *Research in Higher Education, 37*(4), 487-502.

Genereux, R. L., & McLeod, B. A. (1995). Circumstances surrounding cheating: A questionnaire study of college students. *Research in Higher Education, 36*(6), 687-704.

Graham, M. A. (1994). Cheating at small colleges: An examination of student and faculty attitudes and behaviors. *Journal of College Student Development, 35*(4), 255-260.

Granitz, N. A. (2003). Individual, social and organizational sources of sharing and variation in the ethical reasoning of managers. *Journal of Business Ethics, 42*(2), 101-124.

Grijalva, T., Nowell, C., & Kerkvliet, J. (2006). Academic honesty and online courses. *College Student Journal, 40*(1), 180-185.

Gulli, C., Kohler, N., & Patriquin, M. (2007, February 12). The great university cheating scandal. *Maclean's, 120*(5), 32-36. Retrieved from http://www.macleans.ca/homepage/magazine/article.jsp?content=20070209_174847_6984

Haines, V. J., Diekhoff, G. M., LaBeff, E. E., & Clark, R. E. (1986). College cheating: Immaturity, lack of commitment, and the neutralizing attitude. *Research in Higher Education, 25*(4), 342-354.

Hansen, B. (2003). Combating plagiarism: Is the Internet causing more students to plagiarize? *The Congressional Researcher, 13*, 773-796.

Hard, S. F., Conway, J. M., & Moran, A. C. (2006). Faculty and college student beliefs about the frequency of student academic misconduct. *Journal of Higher Education, 77*(6), 1058-1080.

Howard, R. M., & Davies, L. J. (2009). Plagiarism in the Internet age. *Educational Leadership, 66*(6), 64-67.

Hsiao, C., & Yang, C. (2011). The impact of professional unethical beliefs on cheating intention. *Ethics & Behavior, 21*(4), 301-316.

Humbarger, M., & DeVaney, S. (2005). Ethical values in the classroom: How college students responded. *Journal of Family and Consumer Sciences, 97*(3), 40-47.

Jackson, C. J., Levine, S. Z., Furnham, A., and Burr, N. (2006). Predictors of cheating behavior at a university: A lesson from the psychology of work. *Journal of Applied Social Psychology, 32* (5), 1031-1046. doi:10.1111/j.1559-1816.2002.tb00254

Jones, D. R. (2011). Academic dishonesty: Are more students cheating? *Business Communication Quarterly, 74*(2), 141-150.

Jordan, A. (2001). College student cheating: The role of motivation, perceived norms, attitudes, and knowledge of institutional policy. *Ethics & Behavior, 11*(3), 233-247.

Kanfer, R., Wanberg, C. R., & Kantrowitz, T. M. (2001). Job search and employment: A personality-motivational analysis and meta-analytic review. *Journal of Applied Psychology, 86*(5), 837-855.

King, P. M., & Mayhew, M. J. (2002). Moral judgment development in higher education: Insights from the defining issues test. *Journal of Moral Education, 31*(3), 247-270.

Kleiner, C., & Lord, M. (1999, December 14). The cheating games: Everyone's doing it, from grade school to graduate school. *U.S. News & World Report*. Retrieved from http://www.usnews.com/usnews/culture/articles/991122/archive_002427.htm

Koljatic, M. & Silva, M. (2002). Comparison of students' and faculty's perception of occurrence of dishonest academic behaviors. *Psychological Reports, 90* (3), 883-888.

Lanier, M. (2006). Academic integrity and distance learning. *Journal of Criminal Justice Education, 17(2),* 244-261.

Lester, M. C., & Diekhoff, G. M. (2002). A comparison of traditional and internet cheaters. *Journal of College Student Development, 43*(6), 905-911.

Lim, V. K. G., & See, S. K. B. (2001). Attitudes toward, and intentions to report, academic cheating among students in Singapore. *Ethics & Behavior, 11*(3), 261–274.

Luthar, H. K., & Karri, R. (2005). Exposure to ethics education and the perception of linkage between organizational ethical behavior and business outcomes. *Journal of Business Ethics, 61*(4), 353-368.

Ma, H., Wan, G., & Lu, Y. (2008). Digital cheating and plagiarism in schools. *Theory into Practice, 47*(3), 197-203.

McCabe, D. L. (1992). The influence of situational ethics on cheating among college students. *Sociological Inquiry, 62*(3), 365-374.

McCabe, D. L. (1997). Classroom cheating among natural science and engineering majors. *Science and Engineering Ethics, 3*(4), 433-445.

McCabe, D. L. (2001–2002). Cheating: Why students do it and how we can help them stop. *American Educator,* (Winter) 38–43.

McCabe, D. L., Treviño, L. K., & Butterfield, K. D. (2001). Cheating in academic institutions: A decade of research. *Ethics & Behavior, 11*(3), 219–232.

Murdock, T. B., Beauchamp, A. S., & Hinton, A. M. (2008). Predictors of cheating and cheating attributions: Does classroom context influence cheating and blame for cheating? *European Journal of Psychology of Education, 23*(4), 477-492.

Newstead, S. E.; Franklyn-Stokes, A.; and Armstead, P. (1996). Individual differences in student cheating. *Journal of Educational Psychology, 88*(2), 229-241. doi: 10.1037/0022-0663.88.2.229

Nonis, S., & Swift, C. (2001). An examination of the relationship between academic dishonesty and workplace dishonesty: A multicampus investigation. *Journal of Education for Business, 77*(2), 69-77.

O'Fallon, M. J., & Butterfield, K. D. (2005). A review of the empirical ethical decision-making literature: 1996-2003. *Journal of Business Ethics, 59*(4), 375-413.

Okoro, E. A. (2011). Academic integrity and student plagiarism: Guided instructional strategies for business communication assignments. *Business Communication Quarterly, 74*(2), 173-178.

Passow, H. J., Mayhew, M. J., Finelli, C. J., Harding, T. S., & Carpenter, D. D. (2006). Factors influencing engineering students' decisions to cheat by type of assessment. *Research in Higher Education, 47*(6), 643-684.

Pulvers, K., & Diekhoff, G. M. (1999). The relationship between academic dishonesty and college classroom environment. *Research in Higher Education, 40*(4), 487-498.

Rawwas, M.Y.A. and H. Isakson. (2000). Ethics of tomorrow's business managers: The influence of personal beliefs and values, individual characteristics, and situational factors. *Journal of Education for Business, 75*(6), 321-330.

Rettinger, D. A., & Kramer, Y. (2009). Situational and personal causes of student cheating. *Research in Higher Education, 50*(3), 293-313.

Rimer, S. (2003, September 3). A campus fad that's being copied: Internet plagiarism seems on the rise. *The New York Times.* Retrieved from http://www.nytimes.com/2003/09/03/nyregion/a-campus-fad-that-s-being-copied-internet-plagiarism-seems-on-the-rise.html

Roig, M., & Ballew, C. (1992). *Attitudes toward cheating by college students and professors.* Paper presented at the annual meeting of the Eastern Psychological Association (63rd), Boston, MA. Retrieved from ERIC database. (ED349895)

Ruegger, D., & King, E. (1992). A study of the effect of age and gender upon student business ethics. *Journal of Business Ethics, 11*(3) 179-186.

Schrimsher, R. H., Northrup, L. A., & Alverson, S. P. (2009). A survey of Stamford University students regarding plagiarism and academic misconduct. *Plagiary, 3*(2), 1-17.

Shaughnessy, M. F. (1988). The psychology of cheating behavior. Retrieved from ERIC database. (ED349895)
http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED303708

Shepard, J., & Hartenian L. (1991). Egoistic and ethical orientations of university students toward work- related decisions. *Journal of Business Ethics, 10*(4), 303-310.

Shipley, L. J. (2009). Academic and professional dishonesty: Student views of cheating in the classroom and on the job. *Journalism & Mass Communication Educator, 64*(1), 39-53.

Singhal, A. C. (1982). Factors in students' dishonesty. *Psychological Reports, 51* (3), 775-780.

Stevenson, M. J. (1999). Measuring the cognitive moral reasoning of collegiate students-athletes: The development of the Stevenson-Stoll social responsibility questionnaire. *Dissertation Abstracts International: Section B: Sciences and Engineering, 59*(11-B), 6114.

Stuber-McEwen, D., Wiseley, P., & Hoggatt, S. (2009). Point, click, and cheat: Frequency and type of academic dishonesty in the virtual classroom. *Online Journal of Distance Learning Administration, 12*(3), 1-10.

Szabo, A. & Underwood, J. (2003). Academic offences and e-learning: individual propensities in cheating. *British Journal of Educational Technology, 34(4),* 467-477.

Tang, T. L., & Chen, Y. (2008). Intelligence vs. Wisdom: The love of money, Machiavellianism, and unethical behavior across college major and gender. *Journal of Business Ethics, 82*(1), 1-26.

Teodorescu, D. & Andrei, T. (2009). Faculty and peer influences on academic integrity: College cheating in Romania. *Higher Education,* 57, 267-282. Retrieved from http://dx.doi.org/10.1007/s10734-008-9143-3

The Ticker. (2010, November 18). *Cheating on University of Central Florida test was aided by use of textbook questions* [The Chronicle for Higher Education news blog]. Retrieved from http://chronicle.com/blogs/ticker/cheating-on-u-of-central-florida-test-was-aided-byuse-of-textbook-questions/28335

Tucker, A. (2013). Born to be mild. *Smithsonian Magazine, January,* 35-42.

Underwood, J. & Szabo, A. (2004). Academic offences and e-learning: Individual propensities in cheating. *British Journal of Educational Technology, 34,* 467-478.

Vitell, S. J., & Muncy, J. (1992). Consumer ethics: An empirical investigation of factors influencing ethical judgments of the final consumer. *Journal of Business Ethics, 11*(8), 585-597.

Volpe, R., Davidson, L., & Bell. M. C. (2008). Faculty attitudes and behaviors concerning student cheating. *College Student Journal, 42* (1), 164-175.

Watson, G., & Sottile, J. (2010). Cheating in the digital age: Do students cheat more in online courses? *Online Journal of Distance Learning Administration, 13*(1), 1-9. Retrieved from http://www.westga.edu/~distance/ojdla/spring131/watson131.html
Wiley Periodicals. (2014). Student steal, sell test. *Dean and Provost, 15 (6),* 2. doi: 10.1002/dap. website: wileyonlinelibrary.com.

Williams, S. D., & Dewett, T. (2005). Yes you can teach business ethics: A review and research agenda. *Journal of Leadership and Organizational Studies, 12*(2), 109-120.

Whitley, B. (1998). Factors associated with cheating among college students: A review. *Research in Higher Education, 39*(3), 235-27.

Table 1
Factors affecting perceptions of cheating

Factor	I	II	III
I. Having others complete work			
1. You pay for a topic (research) paper from an on-line source and submit it as your own work.	.951		
2. You have your friend complete all of your homework assignments.	.910		
3. You text someone during an exam to get a question answered.	.953		
4. You use a smart phone or other electronic device to search for information during an exam.	.907		
5. You take a picture of an exam and send it electronically to a friend who is taking the exam at another time.	.942		
6. You take a picture of an exam to send to someone who will send the correct answers back.	.949		
7. You have someone else take an online test for you.	.879		
II. Technical issues/reusing old materials			
1. You use a paper you created from a class that you submitted last semester for a class that you are taking this semester but you only make a few changes to the paper.		.646	
2. When taking an on-line exam, you ask your friends for help.		.588	
3. You tell a professor your hard drive/flash drive crashed to get more time on a paper or project.		.631	
4. You tell a professor that technical difficulties prevented you from electronically submitting your work on time when no such problem existed.			

Factor	I	II	III
		.644	
III. Studying from previously created materials			
1. Your roommate ask you to give him your notes from last semester related to the class exams.			
2. You study from a previous (past semester) exam that you acquired from your friend who had the class last year.			.596
3. You review exams taken by friends in previous semester to study.			.878
			.874

Table 2
Average and standard deviation of factors affecting perception of cheating.

Factors	Mean	Standard Deviation
Having others complete work	1.12	.315
Technical issues/reusing old materials	1.32	.309
Studying from previously created materials	1.69	.364

Table 3
Analysis of variance between students and faculty

Statement	df	*F*	Sig.
Your roommate asks you to give him your notes from last semester related to the class exams.	454	1.686	.195
You study from a previous (past semester) exam that you acquired from your friend who had the class last year.	453	.486	.486
You used a paper you created from a class that you submitted last semester for a class that you are taking this semester but you only make a few changes to the paper.	454	54.259	.001
You pay for a topic (research) paper from an on-line source and submit it as your own work.	454	9.073	.003
You work with a group of other students on a research paper but you only do about 2% of the work and tell them to put your name on the paper.	453	1.402	.237
When taking an on-line exam, you ask your friends for help.	454	18.091	.001
You check all the books out of the library related to your research paper so no one else can use that topic.	451	3.778	.053
You pay a person to edit your research paper knowing that grammar accounts for 50% of the assignment grade.	452	9.431	.002
You seduce (for example- by going on a few dates) the “smart” person in class so he/she can help (or complete) most of your assignments in the class.	451	2.701	.101
You have your friend complete all of your homework assignments.	454	7.689	.006
You review exams taken by friends in previous semesters to study.	446	3.918	.048
You text someone during an exam to get a question answered.	452	8.388	.004
You use a smart phone or other electronic device to search for information during an exam.	452	9.319	.002
You take a picture of an exam and send it electronically to a friend who is taking the exam at another time.	454	9.076	.003
You take a picture of an exam to send to someone who will send the correct answers back.	454	8.516	.004
You work with classmates on homework assignments.	451	.407	.524
You tell a professor your hard drive/flash drive crashed to get more time on a paper or project.	451	4.919	.027
You tell a professor that technical difficulties prevented you from electronically submitting your work on time when no such problem existed.	453	6.665	.010
You study using the study guide provided by the professor.	454	8.516	.004
You have someone else take an online test for you.	454	9.265	.002

Analysis of Opportunity to Learn for Students with Disabilities: Effects of Standards-Aligned Instruction

Rolf K. Blank
University of Chicago

John L. Smithson
University of Wisconsin-Madison

Correspondence related to this article should be directed to Dr. Rolf K. Blank, Senior Fellow, Education and Child Development Studies, NORC at the University of Chicago**,** 4350 East-West Highway, Suite 800, Bethesda, MD 20814-4499, blank-rolf@norc.org. The authors wish to acknowledge that research for this paper was supported by a grant from the U.S. Department of Education (#S368A100013) to the Kansas State Dept. of Education. The contents of the paper and presentation do not necessarily represent the policy of the U.S. Department of Education or endorsement by the Federal government. Many thanks to members of the consortium project team who contributed to the research: Rolf Blank, Adam Petermann (CCSSO), John Smithson (WCER), Deb Matthews, Kris Shaw (Kansas), Sandra H. Warren (ASES SCASS adviser), Claire Greer (North Carolina), Wendy Stoica, Andrew Hinkle, Chris Woolard (Ohio), Anne Chartrand, Carolyn Eastman, Brett Moulding, Lani Seikaly (consultants).

Abstract

The paper presents a model for addressing the critical question of opportunity to learn for students with disabilities. The model was tested through a two–year study with schools and teachers in three states. Opportunity to learn analysis is critical in this educational era of push toward access and inclusion. The study results indicate that instruction in grades 4-8 for general education and special education students did not closely align to state content standards. The analysis results indicate that a greater degree of instructional alignment to standards did have a positive impact on student achievement, considering both academic standards and extended standards for students with disabilities. The study findings showed that schools and classrooms providing more inclusive education for students with disabilities had a positive impact on student achievement for all students.

States, local districts, and schools are expected to provide all students with standards-based instruction and inclusive assessments that are well aligned with such instruction. Federal legislation has underscored the right of students with disabilities (SWD) to have access to the general curriculum, instructional content, and tests aligned with standards (IDEA, 1997, 2004; No Child Left Behind Act (NCLB), 2001). We know from annual reports of the U.S. Department of Education that almost all students with disabilities are now participating in annual state student assessments, as specified under NCLB (http://www.ed.gov/esea). The results from recent state assessments show that across all states the average rate of participation was 96 percent of students with disabilities tested in the regular assessment program. However, only 36 percent of these students' scores on the state assessments met their state-defined proficiency level (U.S Office of Special Education Programs, 2013).

The research reported in this paper presents a model for addressing the critical question of opportunity to learn for students with disabilities. The model is tested with results from a two–year research study supported by the U.S. Department of Education conducted with schools and teachers in three states. A priority research question of educators and leaders concerns opportunity to learn, i.e. what is the content of instruction and quality of instructional practices provided in public school classrooms, particularly for students with disabilities in this era of access and inclusion? The research question is operationalized in the present study as: Are students with disabilities, and all students, receiving instruction that is aligned with state standards for learning? And, importantly, what are the differences between the curriculum content and practices taught to students with disabilities as compared to curriculum taught to other students and what are effects on student achievement?

Theoretical Perspective: Research on Opportunity to Learn and Students with Disabilities

Relatively little research has been conducted on the extent to which standards-based instruction at grade level is delivered to students with disabilities, either by general education teachers or special education teachers (Roach, Namisi-Chilungu, et al., 2009). Recent research suggests that students with disabilities in special education classrooms at the same grade level as their general education peers are likely to be getting fewer opportunities to learn expected content (Kurz, Elliott, & Smithson, 2009).

To a large extent, improving instruction and performance of students with disabilities have not been emphasized in standards-based education reform efforts. In the early 2000s, survey research in 34 large school districts found that students with disabilities were not considered in the same way as other students in the context of reforms (Gagnon, McLaughlin, Rhim, & Davis, 2002). Later, Nolet and McLaughlin (2005) summarized their research effort noting that many special educators did not understand the meaning of "curriculum" and saw state content standards and curricular frameworks as too challenging for their students. The study found that many special education teachers reported that it was more important to use instructional time for functional skills than academics; and they showed limited understanding of alternative strategies to meet instructional needs within academically challenging content.

Education policy researchers (e.g., Quenemoen, Thurlow, Moen, Thompson, & Morse, 2003) have noted that students with disabilities have historically had limited access to challenging curriculum, instruction, and assessment. This is sometimes driven by differences in what specific content that access should cover, with some educators believing they need to focus on direct instruction on basic skills and others calling for a full range of rich and challenging grade-level content. Quenemoen et al. (2003) suggest these controversies are intertwined with limited practitioner capacity for effective provision of instructional strategies, interventions, and supports in a standards-based system. Simply put, many special education teachers do not know the content to be taught and many teachers do not know how to teach atypical learners well. Research on opportunity to learn in core academic subjects in general education has developed since the 1990s (Oakes, 1990; Schmidt, et al, 1996). A methodology for use of classroom-based surveys had been tested in several research studies (Porter, 2002; Porter & Smithson, 2001) and evaluations of change in classroom practices were conducted using the survey method and analysis of alignment to state standards (Porter, et al, 2005; Blank, et al, 2006; Smithson &

Blank, 2006). Through collaboration with state specialists, teachers, and researchers and funding support by states and research grants, the Surveys of Enacted Curriculum (SEC) were developed into a web-based system for analyzing and reporting on classroom instructional practices and their relationship to state standards and assessments (Blank, 2010; Blank, et al, 2010). The SEC data tools have been used in over 30 states to analyze math, science and English language arts instruction (see, www.SEConline.org)

In 2010, the Council of Chief State School Officers (CCSSO) responded to the interests of state leaders in special education to develop and submit a successful proposal to the US Department of Education to extend and test the use of the SEC data tools as a model for analyzing opportunity to learn for students with disabilities (Kansas State Department of Education, 2010). The collaborative project led by states and researchers was designed to study opportunity to learn for students with disabilities in comparison to OTL for general education students in the same schools and districts. The project addressed a core need for instruments and data that can assist state and local leaders with a methodology for providing objective evidence of the status of curricular and practice deficits for an at-risk student population and to analyze and report the effects of standards-aligned instruction and opportunity to learn on student achievement. A priority concern voiced by members of CCSSO's State Collaborative on Assessing Students in Special Education identified limited availability of (a) data and appropriate instrumentation for analyzing differences in curriculum and instruction, (b) research-based professional development resources addressing instruction aligned with state standards, and (c) strategies for organizing curriculum and instruction towards improved alignment (see, ASES SCASS state collaborative http://ccsso.org/Resources/Programs/).

The project design included steps to adapt and improve the SEC data collection instruments to address issues of instructional practices, curriculum, and instructional alignment for students with disabilities (CCSSO, 2010a). The data collected through the project were used to analyze the relationship of standards-based instruction to improvement in student achievement. The participating states, districts and schools received assistance in applying their study data in a school-based professional development model to focus instructional improvement strategies on achievement gaps identified through the data analysis. (see project final report, Blank, et al, 2012).

Design and Methodology

The collaborative proposal submitted by the CCSSO research team including state education specialists from states directly participating in the project focused on three research questions that would drive the study design and the data collection and analysis:

1. What is the fidelity of classroom instruction in relation to state adopted content standards and assessments including instruction for students with disabilities and general education students?
2. What are the differences in instructional practices and content taught between special education and general education?
3. What is the effect of instruction students receive to growth in student achievement in mathematics and English language arts and reading?

The project was designed with educators from the three participating states (Kansas, North Carolina, Ohio), special education consultants, and researchers from CCSSO and the Wisconsin Center for Education Research (WCER). The research questions were intended to be addressed through data collected with participating states, districts and schools, as well as to develop, test, and demonstrate the use of research and data tools related to these questions that would be available to the broader community of educators and researchers. The overall project had multiple objectives and several reports and products are available (Blank, et al, 2012, access online through www.SECsurvey.org). The present paper focuses primarily on analysis of data across the sample of teachers and students from all three participating states.

The methodology was based on analysis of instructional practices and content of instruction in a sample of schools and classrooms from three states. The instructional data were analyzed in relation to the content standards for each state, the Common Core State Standards, and academic and extended assessments for each state. The data collection and analysis methodology was based on the Surveys of Enacted Curriculum tools and procedures (Smithson, 2009; Porter, 2002; Blank, et al, 2010). Each participating state was asked to select four to six school districts that had interest in the study research questions and using the data and analyses with their schools. The study targeted grades 4-8 and teachers of ELA and math. Each district was asked to select at least two elementary or middle schools and all teachers in selected grades were asked to participate. This approach ensured that the study sample would be inclusive of teachers with different certifications and assignments for teaching students with disabilities and regular students. The study sample obtained across the three states included 19 districts, 50 schools and 600 teachers (see attached table B for totals by state). The voluntary sample of districts and schools met the study goal of testing differences in instruction between student populations. To address the study objectives, state representative samples of students, teachers and schools were not required.

Data Analysis

Multivariate data analysis was used to address the three research questions. The analysis incorporated three sources of data from participating schools and teachers. Teachers in the 50 participating schools reported on their instructional practices and curriculum content through the SEC online system in spring and fall 2011. Each teacher reported on instruction in English language arts or mathematics (see attached example survey section). School-level program data including teacher assignments, student demographics, and least restrictive environment (LRE) indicator were collected from principals in spring 2011. Student level reading and math scores on state assessments for school years 2009-2010 (prior year) and 2010-2011 (study year), along with information that allowed linking teacher instructional data with their students' test scores. Student scores for the relevant subject (math or language arts) from the prior year served as a prior achievement measure for the multivariate analyses.

Through linking student data to teacher data the analysis could produce more detailed examination of the role that opportunity to learn (OTL) and instructional activities play in the achievement of students. The multivariate regression models were designed to explain differences in student achievement scores controlling for prior achievement and economic

disadvantage and to measure the relative effects of program inclusiveness, opportunity to learn, instructional practices, and students' disability status.

The project data collected through teacher surveys in spring and fall 2011 allowed the project team to report to state leaders and local educators on the degree of fidelity or agreement between the content of instruction provided in classrooms and the standards for student learning required by each of the three project states (Kansas, North Carolina, Ohio). The analysis of alignment is also provided for the statewide assessments used for all students--both general end-of grade academic assessments and modified or alternate assessments used with students with disabilities. Third, we analyze the degree of alignment between current instruction and the Common Core Standards, including fine-grain analysis within topics by grade.

The example graphic displays below show how SEC instructional survey data were reported to schools and teachers as feedback from the study for participants' own use. One firm commitment of the study team to participating districts and schools was assurance that data collected and analyzed through the research design would be available for their own use in analyzing their instruction in relation to state standards. The online SEC data instruments provide data charts that are designed for use by educators to highlight key relationships between study variables. Leader teams from the participating districts and schools received training on analysis and interpretation of their data using charts and graphs similar to these examples. The two SEC data charts show analysis of the relationship between the content of instruction reported by teachers and standards for their state. The "content alignment" analysis is measured through the SEC content framework and application of SEC coding and analysis procedures (Smithson, 2009). The operational definition of alignment in the SEC methodology includes both content topics and level of expectations for student learning (or cognitive demand). Thus for a specific subject and grade level it is possible to analyze the degree of alignment, or consistency, between instruction provided to students and the state standards. (The study reports to educators also provided alignment between instruction and state assessments, and alignment of standards and assessments). The degree of content alignment is reported as a statistic (varying from 0, *no alignment*, to 1, *perfect alignment*), and using the visual displays which allow direct comparison of differences and consistencies between instruction and standards for content topics and expectations for learning. The content analyses of standards and assessments for the three participating states were conducted by subject specialist teams as a part of the research study in June 2011. Content analyses of the Common Core State Standards included in the study were conducted by cross-state specialist teams (CCSSO, 2010b).

The SEC data reporting in Figure 1 shows an example of instructional alignment analysis of English language arts instruction at grade 6 in Kansas classrooms, with comparison to the KS state standards for grade 6. The chart shows data analyzed for 21 grade 6 teachers. (In total, 72 teachers in grades 4-8 in three KS districts reported on instruction in English language arts in the 2011 SEC data collection). The data report informed Kansas educators on the topics and expectations for which instruction differs from standards--for example, the time on instruction is concentrated primarily on the topics Comprehension and Vocabulary while the state standards place high emphasis on Critical Reasoning and Author's Craft. The greatest emphasis in KS standards in the expectations dimension (vertical) is on Analyze/Investigate while the classroom

instructional data shows time emphasis on Memorize/Recall and Perform Procedures. The alignment of the grade 6 classrooms instruction in language arts to the State standards is .37.

The second example of SEC reporting to educators in Figure 2 shows Ohio grade 7 math instructional alignment analysis. Data were reported by 14 grade 7 teachers and the data are compared to OH state math standards. (A total of 87 teachers of math in four Ohio districts participated in the 2011 data collection.) The data chart reveals a heavy emphasis of instruction at grade 7 on Number sense and Operations, while the Ohio standards for grade 7 place more emphasis on Measurement and Basic Algebra. The expectations for learning dimension reported by teachers focus heavily on Perform Procedures, while the Ohio Standards place more emphasis on expectations for Demonstrate understanding, Conjecture/analyze, and Solve non-routine problems. Several math topics that were emphasized in the grade 7 math state standards were reported as having little instructional time– Geometric concepts, Basic algebra concepts, Data displays, and Statistics and Probability. The statistic of alignment across all topics and expectations is .49. The review of the data indicates that misalignment of instruction is largely due to instruction being reported across many topics but instruction is not concentrated in the areas emphasized by state Standards.

Findings for Research Questions

The multivariate analysis results provide findings regarding answers to the research questions. The study teacher-reported data on instruction was linked to the students they taught, and this analytic step provided a method for instructional alignment data to be compared for students with disabilities vs. general education students.

Alignment of instruction to standards and assessments by teacher certification

The data reported in Table 1 shows the degree to which instruction provided by the study teachers was aligned to state standards and assessments, and the data are disaggregated by teachers with special education certification vs. general academic certification. The data on instructional alignment in English Language Arts & Reading (ELAR) show that teachers in the study sample varied substantially according to their certification, regardless of which alignment target is considered. In each analysis of standards and assessments, Special Education teachers reported significantly lower alignment measures compared to their general education peers (see Table 1), and thus students would have fewer opportunities to learn standards-based content. For example, instruction in ELAR by general education teachers is aligned to the state standards for ELAR at the level of .42 (with 1 being perfect alignment), while instruction by special education teachers is aligned at the level of .35 (a significant and substantial difference). This pattern of significant differences persists even for the state extended standards and the modified assessment, where one might expect special educators to place more emphasis than teachers of students in the general population.

While the data for mathematics teachers show no significant differences in instructional alignment between the teacher groups by certification, it is interesting to note the patterns of alignment for the two groups of mathematics. As one might expect, special education teachers

reported higher alignment to the state extended standards as well as to the state modified assessment. Special education teachers also reported slightly higher alignment to state assessments, while general education teachers were slightly more aligned to Common Core Standards. The data suggest that special education teachers may be somewhat more focused on what is assessed than what is in the state standards. Interestingly, teachers in the study, regardless of certification or subject area, tended to report content coverage more aligned to the Common Core state standards than any other instructional target examined. (Note that teachers reported only on their instruction – analysis of alignment was conducted through statistical analysis.) The other notable pattern that emerges in Table 1 is that in general, mathematics teachers tend to report content coverage that is better aligned to each of the instructional targets than reported by language arts and reading teachers.

Differences in instructional practices for students with disabilities vs. general education students

The charts shown in the tables below provide comparisons of instructional activities used with the two categories of student status, and practices are compared for English language arts/reading instruction and Mathematics instruction. Item responses on instructional activities are reported using several scales (e.g., Test preparation, Evaluate argument and evidence, Generate written text, Analyze information, etc.). The study data on instructional activities (classroom practices) aggregated across schools in all three states indicate several key differences by student category that are statistically significant. First, students with disabilities spend less time in language arts instruction engaged in activities focused on Analyzing information and spend significantly less time engaged in Evaluating/critiquing arguments and evidence when compared to their general education peers. The instructional activities focused on Writing and Demonstrate understanding are lower for students with disabilities (although not significant) and Test preparation time is slightly higher for students with disabilities.

The analysis of mathematics instructional activities shows that students with disabilities on average spend more time during mathematics instruction doing math work involving Performing procedures and Taking/preparing for tests than their general education peers. Students with disabilities spend about the same amount of time in activities involving Analyzing information and Demonstrating understanding as the general education students.

Analysis of opportunity to learn and predictors of student achievement

While it is generally accepted that students' opportunity to learn standards-based content and the instructional practices students experience have an impact on student performance, statistical evidence to support these pre-suppositions are not common. Isolated examples of achievement growth and gap reductions can be found for some states and districts, but large scale indicators that capture elements of practice and policy that contribute to explanations of variation in student achievement are rare. The Survey of Enacted Curriculum (SEC) data collection system was selected for this study in order to provide a broad set of indicator measures describing the instruction delivered to general and special student populations. The results serve to inform teachers, administrators, and other educational stakeholders about current practices and provide opportunities for reflection and discussion about appropriate changes to instruction as a result of

these reflections and discussion. In addition, SEC data serve to answer each of the questions posed for the study, whether considering the alignment of instruction to key instructional targets; the instructional practices and content experienced by students with IEP's; the comparison of key general and special education instructional characteristics; or the relationship of these instructional characteristics to student achievement. SEC data provide the relevant indicator measures.

Mathematics achievement. The sample for mathematics achievement analysis comprises 5,004 students across 276 classrooms. Data provided by the states include mathematics and reading achievement scores for students in participating schools for the target year (2011) as well as student achievement data for the prior year (2010). In addition students were flagged on disability status (SWD), and economic disadvantage status (EDS). Identifiers were also provided that permitted students to be associated with the relevant mathematics or language arts teacher to which they were assigned. The data analysis results summarized in Table 3 report findings for multiple indicators of opportunity to learn, as well as several scale measures related to classroom activities. The basic model employed controls for prior achievement, economic disadvantage status, disability status, and the proportion of special education students assigned to category A in the school.

A simple multivariate linear regression model based on these variables yielded an adjusted R^2 of 0.568, with all variables contributing significantly to the model. Adding alignment to the state's content standards increases the adjusted R^2 slightly (to 0.573) and the inclusion of the instructional practice scale measures further increases the adjusted R^2 to 0.587. Thus the addition of these classroom measures provide a modest but positive improvement to the predictive model. While modest, the models indicate that the teacher reports of practice using the SEC instruments do contribute to predicting student achievement, suggesting that the measures have some predictive validity, and in turn increasing confidence in the validity of the teacher self-report data. In order to appreciate the relative impact, the table for Mathematics analysis reports the standardized coefficient for each variable in the model.

The results reported in the table in Table 3 indicate that OTL does have a positive impact on achievement, though at a level somewhat less than the negative effects of economic disadvantage or disability. The model also indicates that schools with higher proportions of students with disabilities spending more time in general education classrooms tend to have higher math achievement scores (i.e. level of inclusion = .048). Each school in the study reported the LRE indicator for the school (LRE=least restrictive environment average percentage of school day with inclusion for SWDs). Among the five scales of instructional practices surveyed, analyzing information represents the one instructional practice that shows a positive impact on student achievement relative to other variables in the analysis.

While the effects are modest, they do indicate that the instruments capture important elements of practice that are linked to achievement, and increase confidence that at the level of school and classroom practice the data has the potential to yield actionable information for teachers that can contribute to increased student performance.

Language arts & reading achievement. The multivariate analysis of student achievement in language arts and reading, comprised 4,004 students in 303 classrooms in the sample from participating schools across three states. The results of the analysis are shown in Chart 4. The analysis results for ELAR achievement look quite different than the results reported for mathematics. In general, multivariate regression models in language arts tend to report higher adjusted R^2, however this is largely due to the greater predictive power of prior achievement in language arts (adj. R^2= .70 versus .58 for math). While classroom practices and standards alignment measures do provide statistically significant contributions to the predictive model (e.g., Generate written text = .22), the direction of the influence varies from one indicator to another and in ways that may appear to be counter-intuitive (alignment to state standards = -.178). This may in part be due to state achievement tests being primarily tests of student reading scores. The results do support the assertion that opportunity to learn and pedagogical indicators can contribute to explaining variations in student achievement gains and thus inform curriculum decisions designed to optimize student performance. The model also indicates that schools with higher proportions of students with disabilities spending more time in general education classrooms tend to have higher achievement scores (i.e. level of inclusion = .048)

The adjusted R^2 for the equation represented in the language arts/reading Table is .701. The results for ELAR analysis represent a better model fit than the results for mathematics. Among the classroom practice measures, generating written text tends to be the best predictor of achievement. Each of the measures of opportunity to learn have a positive impact on achievement and the combined Standardized Beta Coefficients exceed the negative effects of economic disadvantage and disability status.

Conclusions

The analysis results from the study of opportunity to learn in a sample of classrooms, schools and districts across three states provide several types of important evidence that contribute to understanding of the relationship of instruction to student achievement. First, the data analysis identified where and how instructional practices and content of instruction in the 50 sample schools differed from the standards established by their states. Across the focus grades 4-8 for this study, the instruction reported by teachers for general education and special education students did not closely align to state content standards, both in distribution of instructional time by topic and in the expectations for learning that are emphasized. However, the analysis did show that a greater degree of instructional alignment to standards did have a positive impact on student achievement. And, the positive relationship of alignment to standards to achievement held for both regular academic standards and extended standards for students with disabilities. Second, evidence was provided regarding questions raised by special education experts about the extent to which students with disabilities are receiving a standards-based education. Schools and classrooms providing more inclusive education for students with disabilities had a positive impact on student achievement for all students. However, overall, students with disabilities had average achievement scores that were significantly lower than general education students. The study data also identified several areas of instructional practices in which students with disabilities receive different levels of instructional time and emphasis than general education students, including less time on writing, analysis of information, and evaluating evidence and arguments, and more time on test preparation.

The study results also demonstrated the use of the Surveys of Enacted Curriculum data instruments for use in special education research and analysis of differences in opportunity to learn. The study had a practical benefit for participants in that school teams participating in the study received school reports with graphic displays that provided a baseline picture of instructional practices in ELA and mathematics in classrooms. The data charts were used to analyze instruction by classroom student composition, teacher preparation, and variation instructional content by grade in relation to prior state standards as well as the new Common Core Standards that were being introduced as the study took place. Further extension of the research is possible since the research model and data tools can be used to track changes in instructional practices over time, and analyze effects of instructional improvement initiatives. The multivariate data analyses indicated that opportunity to learn, classroom activities and inclusion policies all contribute to student performance to some degree. The cross-state data provide a descriptive baseline while suggesting dynamics and relationships that deserve further investigation. A basic question underlying all of the results is the degree to which the findings from this study are generalizable. Considering the diversity of teachers and programs represented in the data-set, collected from approximately 300 teachers in each subject across the three states, there is good reason to believe the results are reflective of the conditions for teaching in special education more broadly, and the relationship of instruction to student achievement. However results from further studies and other data collection efforts are needed to either confirm or alter the picture of special education portrayed in this report.

Discussion

The research analysis from this study of opportunity to learn based on analysis of instructional alignment to standards provides results that can be generalized to the larger population of teachers and students, potentially providing insights into basic elements of mathematics and reading instruction that have relevance for teachers, administrators and researchers beyond the boundaries of the schools, districts and states that participated in the study. Through prior collaboration with many state education specialists and teachers, the CCSSO researchers and state education leaders had identified a specific need for instruments and data that can assist state and local leaders with objective evidence of the status of curricular and practice deficits for students with disabilities.

With the adoption of the Common Core Standards by many states, leaders identified the need to improve methods of professional development for local leaders and teachers that would highlight the key transitions in instruction needed for implementation of the Common Core. They also sought research-based evidence of the gaps in student achievement and enacted curriculum in classrooms that would drive the argument for improving practices through professional development with all teachers. The model provided by this study can now be used by state and local education leaders to advance their work to align instruction and curriculum consistent with the Common Core Standards. The tools for measuring and reporting on the variation in instruction aligned with standards demonstrated by the analysis model will serve educators and researchers in further efforts to align instruction to standards. When combined with student achievement data that can be associated with specific teachers reporting their practice using the SEC instruments, the SEC data-set provides a unique opportunity to examine the predictive

properties for a variety of OTL and classroom activity measures in explaining variation in student achievement scores. One of the largest successes of the study then has been acquiring access to student achievement data from participating states for the schools in the study in a manner that permitted making the connection of SEC teacher reports with performance data for the students in their class during the time of the study.

The evidence from this study also highlight the need to carefully analyze and specify the differences in instructional practices and content being delivered to students with disabilities as compared to instruction provided for general education students. The evidence from this study show that the specific differences are related to differences in tested achievement outcomes.

References

Blank, R., Smithson, J., Porter, A., Nunnaley, D., & Osthoff, E. (2006) Improving instruction through
school wide professional development: Effects of the data-on-enacted-curriculum model, *ERS Spectrum*, Spring.

Blank, R., (2010) Surveys of Enacted Curriculum and the Council of Chief State School Officers Collaborative, in *Research and Practice Pathways in Mathematics Education*, Tate, King, Anderson (eds), 2011, Reston, VA: NCTM, pp.21-32.

Blank, R., Smithson, J., Deeter, T., (2010) *Improving methods of aligning instruction to standards and*
assessments for English language learners and analyzing the relationship of alignment to student achievement. Final report: Iowa SEC ELL State Consortium Project, Study Conducted under a grant from U.S. Department of Education. CCSSO, Washington, DC.

Blank, R., Matthews, D., Smithson, J. (2012) *Closing the Opportunity Gap for Students with Disabilities:*
Analyzing Alignment of Instruction and Standards in English Language Arts and Mathematics, Summary Report. Kansas EAG State Consortium SEC Special Education Project (US ED grant #S368A100013) Council of Chief State School Officers, Washington, DC, December 2012.
http://www.ccsso.org/Resources/Programs/Surveys_of_Enacted_Curriculum_(SEC).html go to SECPD Online.

CCSSO (2010a) Project design prospectus: Develop instrumentation to analyze fidelity of instruction for students with disabilities in relation to standards and assessments and report on opportunity to learn and student achievement. Council of Chief State School Officers, Washington, DC, November 2010.

CCSSO (2010b) Description of SEC Content Analysis of Common Core State Standards and Initial Data Analysis Results, CCSSO SEC State Collaborative, Washington, DC.

Gagnon, J. C., McLaughlin, M. J., Rhim, L. M., & Davis, G. A. (2002). Standards-driven reform policies at the local level: Report on a survey of local special education directors in large districts. *Journal of Special Education Leadership, 15*(1), 3-9.

Kansas State Department of Education (2010) Project proposal: Develop instrumentation to analyze fidelity of instruction for students with disabilities in relation to standards and assessments and report on opportunity to learn and student achievement. Proposal for Enhanced Assessment Grant to U.S. Department of Education, May 2010.

Kurz, A., Elliott, S. N., & Smithson, J. L. (2009). *Alignment of the intended, planned, and enacted curriculum in general and special education and its relation to student achievement.* Manuscript submitted for publication.

Nolet, V. & McLaughlin, M. J. (2005). *Accessing the general curriculum: Including students with disabilities in standards-based reform (2nd ed.).* Thousand Oaks, CA: Corwin Press.

Oakes, J. (1990). *Multiplying inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and science*. Santa Monica, CA: Rand.

Porter, A.C. & Smithson, J. (2001). Are content standards being implemented in the classroom? A methodology and some tentative answers. In S.H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states.* Chicago: National Society for the Study of Education.

Porter, A. C. (2002, October). Measuring the content of instruction: Uses in research and practice. Edu*cational Researcher 31*(7), 3-14.

Porter, A. C., Blank, R. K., Smithson, J. L., & Osthoff, E. (2005). Place-based randomized trials to test
the effects on instructional practices of a mathematics/science professional development program for teachers. *The Annals of the American Academy of Political and Social Science, 599*(1), 147–175.

Quenemoen, R., Thurlow, M., Moen, R., Thompson, S., & Morse, A. B. (2003). Progress monitoring in an inclusive standards-based assessment and accountability system (Synthesis Report 53). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Roach, A. R., Namisi-Chilungu, E., LaSalle, T. P., Talapatra, D., Vignieri, M. J., & Kurz, A. (2009). *Alternate assessment based on modified achievement standards: Opportunities and options for facilitating and evaluating access to the general curriculum for students with disabilities*. Manuscript submitted for publication.

Schmidt, W.H., McKnight, C.C., Valverde, G.A., Houang, R.T., & Wiley, D.E. (1996). *Many visions, many aims: A cross-national investigation of curricular intentions in school mathematics.* Analysis of TIMSS Data. Boston: Kluwer Academic Publishers.

Smithson, J., & Blank, R. (2006). Indicators of quality of teacher professional development and

instructional change using data from surveys of enacted curriculum: Findings from NSF MSP-RETA project. Washington, DC: Council of Chief State School Officers. Retrieved from http://jsmithson. wceruw.org/reference/indicators_tchrpd_smithson-blank.pdf

Smithson, J.L. (2009) Coding procedures for curriculum content analysis, WCER, University of Wisconsin-Madison. www.SEConline.org

Figure 1

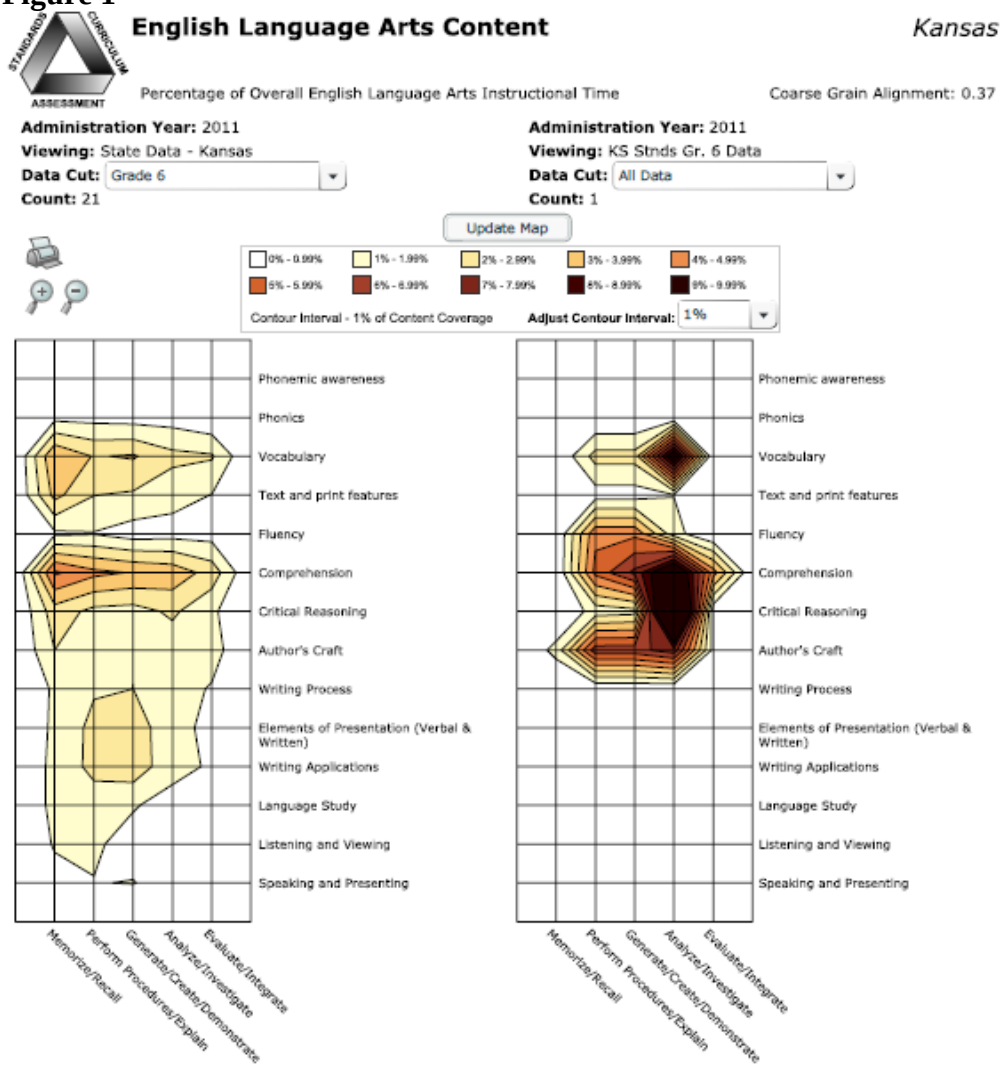


Figure 2

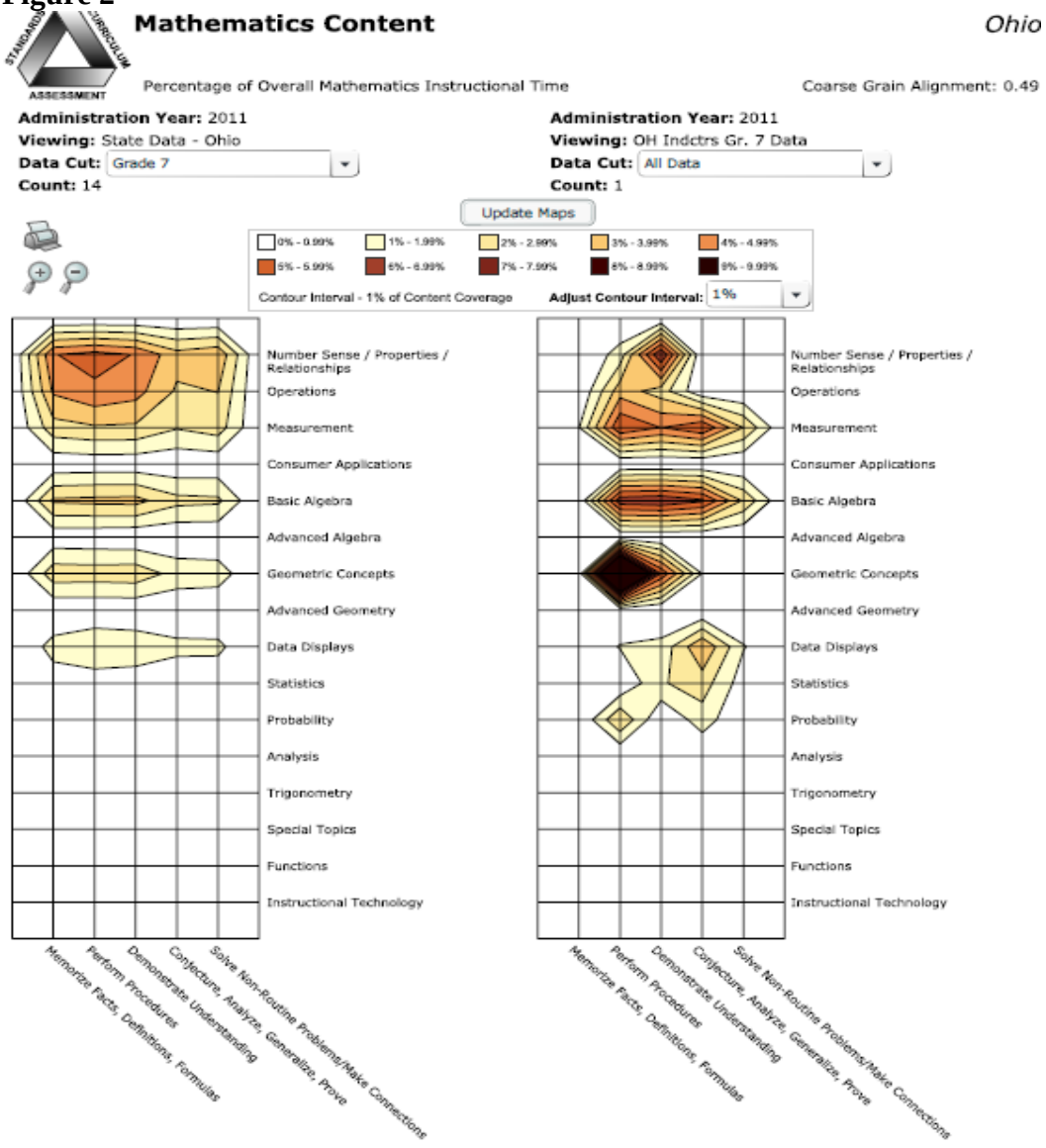
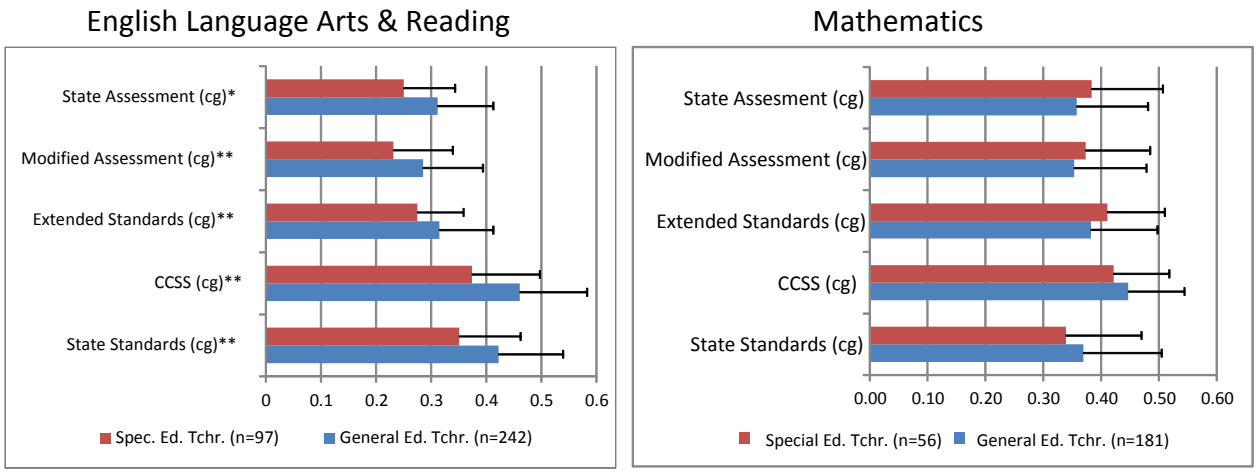
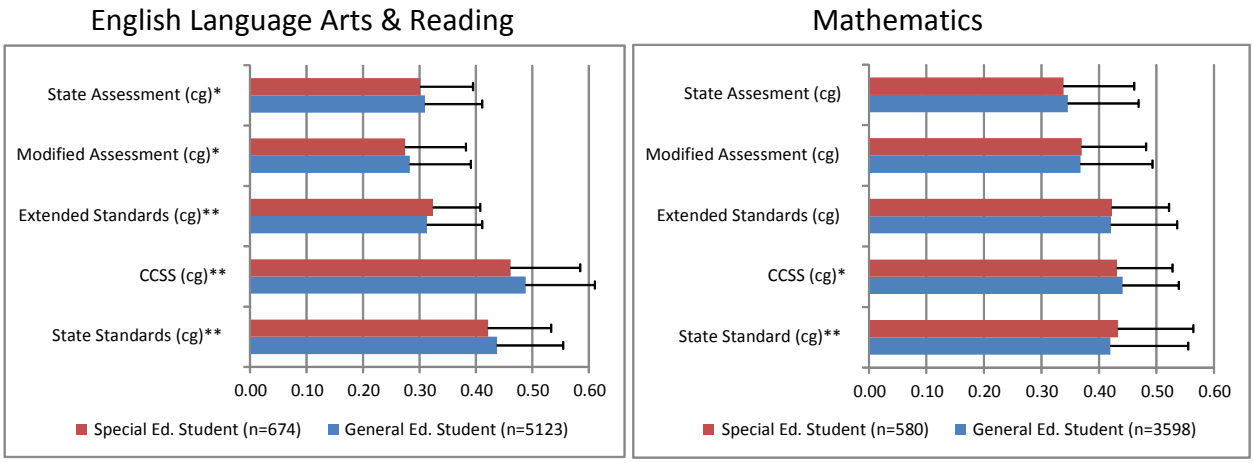

Mathematics Content
Ohio
Percentage of Overall Mathematics Instructional Time
Coarse Grain Alignment: 0.49
Administration Year: 2011
Viewing: State Data - Ohio
Data Cut: Grade 7
Count: 14
Administration Year: 2011
Viewing: OH Indctrs Gr. 7 Data
Data Cut: All Data
Count: 1
Update Maps
0% - 0.99%
1% - 1.99%
2% - 2.99%
3% - 3.99%
4% - 4.99%
5% - 5.99%
6% - 6.99%
7% - 7.99%
8% - 8.99%
9% - 9.99%
Contour Interval - 1% of Content Coverage
Adjust Contour Interval: 1%
Number Sense / Properties / Relationships
Operations
Measurement
Consumer Applications
Basic Algebra
Advanced Algebra
Geometric Concepts
Advanced Geometry
Data Displays
Statistics
Probability
Analysis
Trigonometry
Special Topics
Functions
Instructional Technology
Memorize Facts, Definitions, Formulas
Perform Procedures
Demonstrate Understanding
Conjecture, Analyze, Generalize, Prove
Solve Non-Routine Problems/Make Connections

Table 1
Instructional Alignment to Standards and Assessments by Teacher Certification



** Significant at p < 0.05* *** Significant at p < 0.01* *Whisker = 1 Standard Deviation*

Table 2
Instructional Alignment by Students' Disability Status



** Significant at p < 0.05* *** Significant at p < 0.01* *Whisker = 1 Standard Deviation*

Table 3
Mathematics Multivariate Regression Equation Explaining Student Achievement Scores

Variable	Standardized Beta Coefficient	Significance
Prior achievement	0.70	0.000
Disability status	-0.096	0.000
Economic Disadvantage	-0.069	0.000
Level of SWD inclusion	0.048	0.001
State Standard (pre-CCSSM)	0.065	0.000
CCSSM	-0.073	0.000
NCTM	0.132	0.000
Analyzing Information	0.040	0.000

English Language Arts & Reading regression equation explaining student achievement scores

Variable	Standardized Beta Coefficient	Significance
Prior Achievement	.777	.000
Disability Status	-.076	.000
Economic Disadvantage	-.065	.000
Level of SWD inclusion	.045	.000
State Standard	-.178	.000
CCSS	.090	.000
State Mod. Test	.062	.000
State Ext. Standard	.030	.004
Generate Written Text	.022	.037

Appendix A

Section of SEC teacher survey—Mathematics instructional practices

AMOUNT OF INSTRUCTIONAL TIME

0 - None
1 - Little *(Less than 10% of instructional time for the school year)*
2 - Some *(10-25% of instructional time for the school year)*
3 - Moderate *(26-50% of instructional time for the school year)*
4 - Considerable *(More than 50% of instructional time for the school year)*

How much of the mathematics instructional time in the target class do students use to engage in the following tasks?		*None*	*Little*	*Some*	*Moderate*	*Considerable*
IP.1	Listen to the teacher explain, or observe the teacher demonstrate or model a math procedure or solve a problem	⓪	①	②	③	④
IP.2	Read and comprehend mathematics information from multiple sources	⓪	①	②	③	④
IP.3	Collect, summarize, and/or analyze information or data from multiple sources	⓪	①	②	③	④
IP.4	Present or demonstrate to others	⓪	①	②	③	④
IP.5	Work *individually* on mathematics assignments	⓪	①	②	③	④
IP.6	Participate in whole-class discussions about mathematics	⓪	①	②	③	④
IP.7	Engage in a writing process to support arguments with evidence	⓪	①	②	③	④
IP.8	Use hands-on materials	⓪	①	②	③	④
IP.9	Work in pairs or small groups on mathematics exercises, problems, investigations, or tasks	⓪	①	②	③	④
IP.10	Engage in learning activities outside the classroom	⓪	①	②	③	④
IP.11	Use computers, calculators, or other technology to learn, practice or explore mathematics	⓪	①	②	③	④
IP.12	Maintain and reflect on a portfolio of their own work	⓪	①	②	③	④
IP.13	Practice test-taking strategies	⓪	①	②	③	④
IP.14	Take a quiz or test	⓪	①	②	③	④

Appendix B

KS EAG Consortium Project: Study Sample by State

	Districts	Schools	Teachers		
			ELA	Math	SwSCD
Kansas	3	15	63	72	11
North Carolina	5	16	86	88	23
Ohio	4	19	87	115	10

Note: The table reports the number of school districts, schools, and teachers participating in the study during 2011 to 2012. Teacher sample numbers indicate the number of teachers in grades 4-8 completing the SEC teacher survey on classroom instructional practices. Teacher survey categories = English language arts, Mathematics, and Students with significant cognitive disabilities.

Trend Analysis of Educational Investments and Outcomes

Robert J. Griffore, Ph.D.
Michigan State University
griffore@msu.edu

Lillian A. Phenice, Ph.D.
Michigan State University

Meng-Chuan Hsieh, M.A.
Michigan State University

Correspondence related to this article should be directed to Dr. Robert J. Griffore, Department of Human Development and Family Studies, 552 W. Circle Drive, 13H Human Ecology Building, Michigan State University, East Lansing, Michigan 48824 (griffore@msu.edu).

Abstract

It is generally expected that funds invested in public instruction will strengthen pupil achievement. Similarly, it is expected that enrollments of young children in preschool education should lead to higher achievement. In a trend analysis, we examined three variables as predictors of grade 8 achievement in 2007: (1) state expenditures per student in 2000, (2) state level percentage of 3 year old and 4 year old children who attended school in 2000, and (3) state level grade 4 achievement scores in 2003. State expenditures per pupil and percent participation in early childhood education at the state level were not significant predictors of grade 8 achievement. Grade 4 achievement explained more than 80 percent of the variance in grade 8 achievement.

On August 7, 2013, Bill Gates made a presentation at the National Governors Association Conference in which he stated what has been known for several years. We have been investing more in public education, measured as per-student achievement, while educational achievement has been flat (Gates, B., 2013). His presentation is based on a graph in which the curve of investment in public education is shown to steadily increase from 1975 to 2007, while both reading and math achievement remained virtually unchanged during the same time period.

Educational Investments

One would expect intuitively that greater investments in education would lead to improvements of several kinds. More investment in the form of bonds and sinking funds can be expected to support building and renovation. Investments in supplies and equipment will support classroom instruction. Educational theorists, advocates, and researchers, would likely postulate that children who are beneficiaries of enriched educational resources will experience greater school success. Thus it is not unreasonable to assume that school achievement should increase as when there are increases in per-pupil expenditures.

The assumption that it is wise to invest in public education is not completely without a base of evidence; however the base of evidence is not completely clear and is somewhat confusing. For example, there is evidence that school spending is associated with school success (Greenwald, Hedges & Laine, 2006; Verstegen & King, 1998). However, there is evidence as well of an inverse association between funding and achievement. Sharp (1993) found a slight negative correlation between spending and achievement at all grades except grade 11 and in all subject matter areas. Other research has suggested that the association of spending with achievement is uncertain (Ludwig & Bassi, 1999).

It is assumed that early childhood education, especially in the form of a well designed and expertly delivered curriculum, can compensate for a variety of deficits and risks that prevent low school achievement. The literature indicates that providing educational resources in the form of effective early childhood education programs can make a difference in achievement in the primary and later grades. This literature can be considered thoroughly established and well documented. These findings have been documented in numerous recent summaries of the literature (Barnett, 2011; Campbell & Ramey, 1994; Magnuson & Waldfogel, 2005; Mervis, 2011).

However, it is important to note that studies of the effects of early childhood education tend to describe the outcomes of unique educational programs, which are designed to achieve specific purposes for particular children. Much of the confidence placed in positive outcomes of early childhood education comes from experimental programs from the 1960s and 1970s (Besharov & Ramey, 2010). Early childhood education outcomes are related to the specific purposes for which programs are designed (Hines, McCartney, Mervis & Wible, 2011). It is important to acknowledge the impact of particular program models (Barnett, 2011). Some programs are designed to improve specific outcomes, such as the executive function in children four to twelve years old (Diamond & Lee, 2011). Clements and Sarama (2011) discuss teaching for early mathematics competence and later mathematics achievement.

There is no doubt that early childhood education can lead to higher achievement for particular children who experience a specific form of curriculum intervention in a particular context. However, is it to be expected that all early childhood education will boost later achievement? A series of large-scale experimental studies, perhaps varying both curriculum approaches and distinct groups of participating children, might demonstrate what has already been thoroughly documented – that when particular curriculum interventions are matched with particular groups of children, the outcomes are predictably good. This approach would not resolve the question of whether the level of early childhood education investment at the state level will predict educational outcomes.

State Level Analysis of Educational Investments and Outcomes

There is much variation across states, both in the provision of educational resources to children and in educational outcomes. Considerable attention is paid to describing state-level differences in educational inputs (National Center for Educational Statistics, 2004) and outputs (National

Center for Educational Statistics, 2012). Tabular presentation of input and output data encourage ranking of state with regard to a variety of indicators and metrics. The National Education Association ranks states on many indicators (National Education Association, 2013), as does the American Legislative Exchange Council (2013). *Education Week* (2013) announced the availability of "Quality Counts", a report on state educational policy and outcomes.

When states are ranked, profiles of strengths and deficits become more evident. Presumably policy makers in states that are ranked low will be motivated to address ways of optimizing educational issues and seek higher rakings. They may decide to make more substantial investments in educational inputs with the goal of boosting educational achievement. The problem is, as Mr. Gates has observed, the curve of educational achievement may continue to be flat.

Method

The analysis presented in this paper provides evidence in support of Mr. Gates' assertion about the relationship between educational investment and outcomes. It also expands the analysis to include early childhood education, which is a form of investment. A trend analysis is used in this study. While trend studies can be limited due to the use of data from samples at different points in time (Wimmer &Dominick, 2011), the method can be useful for considering patterns of overall stability and change over time across different samples. Trend analysis is used in epidemiologic research (Rosenberg, 1997). It can also be useful for educational policy research for identifying existing patterns. It can provide a broad perspective of aggregate demographic and geographic data over a period of several years. This is not possible through the use of case studies, which involve special selection of particular child participants and/or are based on particular expressions of curriculum concepts, thus limiting generalizability. Thus, in this analysis we examine whether per pupil expenditures or participation in early childhood education predict eighth grade academic achievement, along with a third predictor – achievement at the elementary level.

This analysis makes use of existing public state-level summary data for 50 states. The state is the unit of analysis. Individual pupil data were not analyzed. The following variables were used in this analysis.

- Percentage of 3 year old and 4 year old children in 2000 who attended school (United States Census Bureau, 2000)
- State per pupil expenditure in 2000 (National Center for Education Statistics, 2012a).
- State level average National Assessment of Educational Progress (NAEP) average grade 4 reading scale scores in 2003 (National Center for Educational Statistics, 2012b)
- State level average NAEP grade 4 mathematics scale scores in 2003 (National Center for Educational Statistics, 2012b)
- State level average NAEP grade 8 reading scale scores in 2007 (National Center for Educational Statistics, 2012b)
- State level average NAEP grade 8 mathematics scale scores in 2007 (National Center for Educational Statistics, 2012b).

Results

Descriptive statistics were computed, and two regression models were examined: a model of reading achievement and a model of mathematics achievement. SPSS Version 21 was used to examine reading and math models.

- *Per-Pupil Expenditures.* There was substantial variability in state level per pupil expenditures for instruction in the year 2000. The range was $4,378 to $10,337 per pupil. The mean was $6,779 (SD = $1,311).
- *Percentage of 3- and 4-Year Olds Enrolled in School*. Likewise, there was considerable variation in state level percentage of 3 and 4 year olds enrolled in school in the year 2000. The lowest percentage was 34.4, and the highest was 63.2. The mean percentage was 47.5 (SD= 6.73).
- *NAEP 2003 Grade 4 Reading.* The range of state level NAEP grade 4 reading achievement in year 2003 was from 203.0 to 228.0. The mean was 218.02 (SD = 6.45).
- *NAEP 2003 Grade 4 Math.* The range of state level NAEP grade 4 math achievement in year 2003 was from 223.0 to 243.0. The mean was 234.5 (SD = 5.41).
- *NAEP 2007 Grade 8 Reading.* The range of state level NAEP grade 8 reading achievement in year 2007 was from 250.0 to 273.0. The mean was 262.48 (SD = 6.27).
- *NAEP 2007 Grade 8 Math*. The range of state level NAEP grade 8 math achievement in year 2007 was from 265.0 to 298.0. The mean was 281.34 (SD = 7.47).

Correlations were computed. There was a relatively low correlation between per-pupil expenditure and grade 8 math ($r = .31$, $n = 50$, $p = <.05$). Likewise, there was a similarly low correlation between per-pupil expenditures and grade 8 reading ($r = .35$, $n = 50$, $p = <.05$). The correlation of enrollment of young children in school and grade 8 math is not significant ($r = .06$, $n = 50$, $p = >.05$). Likewise, the correlation between enrollment of young children in school and grade 8 reading is not significant ($r = .11$, $n = 50$, $p = >.05$).

Closer examination of the data provides insights into these low correlations. For example, New Jersey spent $10,337 per pupil in 2000 and had an average NAEP Grade 8 math score in 2008 of 292. New Jersey spent $5,667 per pupil in 2000 and had an average NAEP Grade 8 math score in 2008 of 289. Thus a difference of almost $5,000 resulted in a difference of only 3 points. And for New Jersey's investment of $10,337 per pupil in 2000, the NAEP Grade 8 reading score in 2008 was 270. For an investment of $5,632 per pupil in 2000, South Dakota also had an average NAEP Grade 8 reading score in 2008 of 270. Thus a difference of almost $5,000 resulted in no difference in NAEP reading score.

It is useful to consider also the following examples regarding early childhood education. In North Dakota, in 2000 only 34.4 percent of young children were in school. The average NEAP math score was 292. New Jersey had 63.2 percent of young children in school in 2000, yet in 2007 the New Jersey NAEP Grade 8 math average score was 289 – lower than North Dakota. Similarly, In South Dakota in 2000, only 39.9 percent of young children were in school. The

average NEAP reading score was 270. While New Jersey had 63.2 percent of young children in school in 2000, in 2007 the New Jersey NAEP Grade 8 reading average score was 270 – the same as South Dakota.

A multiple regression analysis was conducted to predict NAEP 2007 grade 8 reading scores from the following variables: (1) per pupil expenditures in 2000, (2) percentage of 3 year old and 4 year old children who attended school in 2000, and (3) NAEP average grade 4 reading scale scores in 2003. The regression was significant ($F (3, 46) = 74.88$, $p < .001$, $R^2 = .82$). Of the three predictor variables, only NAEP average grade 4 reading scale scores in 2003 was significant ($\beta = .91$, $t (49) = 13.78$, $p < .001$). State expenditure per pupil in 2000 was not a significant predictor of reading achievement ($\beta = .02$, $t (49) = .38$, $p > .05$. State level percent of 3 and 4 year olds enrolled in school in 2000 was not a significant predictor of reading achievement ($\beta = -.08$, $t (49) = -1.30$, $p > .05$).

A multiple regression analysis was conducted to predict NAEP 2007 grade 8 math scores from the following variables: (1) per pupil expenditures in 2000, (2) percentage of 3 year old and 4 year old children who attended school in 2000, and (3) NAEP average grade 4 math scale scores in 2003. The regression was significant ($F (3, 46) = 69.82$, $p < .001$, $R^2 = .82$). Of the three predictor variables, only NAEP average grade 4 math scale scores in 2003 was significant ($\beta = .89$, $t (49) = 13.79$, $p < .001$). State expenditure per pupil 2000-2001 was not a significant predictor of math achievement ($\beta = .05$, $t (49) = .38$, $p > .05$). State level percent of 3 and 4 year olds enrolled in school in 2000 was not a significant predictor of math achievement ($\beta = -.07$, $t (49) = -1.30$, $p > .05$).

Discussion

Using state-level data, this analysis explored the extent to which state-level participation in early childhood education, statewide expenditures for public instruction, and subsequent achievement in grade 4 predicted later academic achievement in grade 8. The results, based on state-level data, indicate that the best predictor of achievement in grade 8 was achievement in grade 4. In the model examined here, per-pupil expenditures were not a significant predictor of achievement. Nor was participation in early childhood education a significant predictors of achievement.

The results of this analysis should not be interpreted to indicate that it would be appropriate to reduce per-pupil expenditures or support for early childhood education for young children. However, it is important to examine educational investments in perspective. It is not clear that achievement within a state will be increased simply by increasing per-pupil expenditures or by augmenting support for early childhood education. Without specifically addressing how these resources are delivered based on particular contexts and including individual child needs, there will likely be little improvement in educational outcomes.

When districts receive funds, there is no assurance that the funds will be spent in ways that are matched with children's needs and abilities or with regard to the capacity of the district to deliver instruction. Funds may not be targeted toward the educational needs of children or congruent

with the realities of instructional possibilities. Therefore, investments of dollars will generally have unpredictable outcomes.

When educational resources are strategically targeted toward boosting school achievement, higher achievement will likely follow. Presently, resources may not be reliably matched with the needs or abilities of individual children within districts. Educational policy makers could benefit from information proposed provided by school districts concerning optimal allocation of resources, with a specific focus on how funds will be allocated for specific curriculum areas, such as reading, math, or science.

In addition, the most advantageous point of intervention may not necessarily be pre-kindergarten education. Pre-kindergarten education is necessary for the socialization of young children for good citizenship. It is evident that delivery of particular early education programs to select groups of young children is successful. However, the efficacy of early childhood education depends on delivery of particular kinds of early childhood education to children who need these particular experiences. The specifics of process are crucial to success, in that same way that specific processes of investment of dollars in public instruction shapes and determines educational outcomes.

In this analysis, achievement in grade 4 was the best predictor of achievement in grade 8. This suggests the efficacy of strengthening educational quality and outcomes in grades 1, 2, and 3. Funding could be targeted to districts that show poor trend patterns in the early grades, specifically targeted toward curriculum improvement, appropriate human resources or other district needs. The result of the present analysis suggests that optimizing allocations leading to and including the fourth grade level can be advantageous for predicting achievement in later years.

References

American Legislative Exchange Council (2013). Report Card on American Education. Accessed August 4, 2013 at http://www.alec.org/publications/report-card-on-american-education/

Barnett, W.S. (2011). Effectiveness of early educational intervention. *Science, 333 (6045*), 975-978.

Besharov, D.J. & Ramey, C. (2008). Preschool puzzle. *Educationnext.* Retrieved from http://educationnext.org/preschool-puzzle/

Campbell, F.A. & Ramey, C. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low income families. *Child Development, 65,* 684-698.

Clements, D.H. & Sarama, J. (2011). Early childhood mathematics intervention. *Science, 333 (6045)*, 968-970.

Diamond, A. & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science, 333 (6045*), 959-964.

Education Week (2013). States Show Spotty Progress on Education Gauges. Accessed August 4, 2013 at http://www.edweek.org/ew/articles/2013/01/10/16sos.h32.html?intc=EW-QC13-TOC

Gates, B. (2013). Flip the curve: Student achievement vs. school budgets. Accessed on August 8, 2013 at http://www.huffingtonpost.com/bill-gates/bill-gates-school-performance_b_829771.html

Greenwald, R., Hedges, L, & Laine, R. (1996). The effect of school resources on student achievement." *Review of Educational Research, 66,* 361-396. doi: http://dx.doi.org/10.2307/1170528

Hines, P., McCartney, M., Mervis, J. & Wible, B. (2011). Laying the foundation for a lifetime learning. *Science, 333 (6045),* 951.

Ludwig, J. & Bassi. L. (1999). The puzzling case of school resources and student achievement. *Educational Evaluation and Policy Analysis, 21 (4),* 385-403.

Magnuson, K. & Waldfogel. J. (2005). Early childhood care and education: Effects on ethnic and racial gaps in school readiness. *The Future of Children, 15 (1*), 169-196.
Mervis, J. (2011). Past successes shape effort to expand early childhood education. *Science, 333 (6045),* 952-956.

National Center for Education Statistics (2012a). NCES Common Core of Data. Table 8-11. Current expenditures per pupil for elementary and secondary public schools, by state: 2000, 2005, and 2009. Accessed August 4, 2013 at http://www.nsf.gov/statistics/seind12/c8/c8s1o11.htm

National Center for Education Statistics (2012b). NAEP State Comparisons. U.S. Department of Education. Institute of Education Sciences.
http://nces.ed.gov/nationsreportcard/statecomparisons/

National Education Association (2013). Rankings of states and estimates of school statistics. Access August 4, 2013 at http://www.nea.org/home/44479.htm

Rosenberg, D. (1997). Trend analysis and interpretation. Division of Science, Education and Analysis, Maternal and Child Health Bureau, Department of Health and Human Services. http://mchb.hrsa.gov/publications/pdfs/trendanaylsis.pdf

Sharp, W. (1993). School spending: Is there a relationship between spending and student achievement? A Correlation study of Illinois schools. Paper presented at the annual meeting of the American Education Finance Association, Albuquerque, NM, March.

United States Census Bureau (2000). Census Summary File 3, Table PCT 23 and the National Center for Educational Statistics (United States Department of Education) 1999 Household Education Survey. http://nieer.org/resources/facts/index.php?FastFactID=10

Verstegen, D.A. & King, R.A (1998). The relationship between school spending and student achievement: A review and analysis of 35 years of production function research. http://ezproxy.msu.edu/login?url=http://search.proquest.com/docview/62451622?accountid=12598

Wimmer, R.D. & Dominick., J.R. (2011). *Mass media research: An introduction*. Boston: Wadsworth.