

# MID-WESTERN EDUCATIONAL RESEARCHER

Official Publication of the Mid-Western Educational Research Association

Volume 21, No. 4—Fall 2008



University of West Georgia



[www.mwera.org](http://www.mwera.org)

### *On the Cover*

The University of West Georgia is a rapidly growing institution located approximately 50 minutes west of Atlanta, Georgia. The University enrolls over 11,000 graduate and undergraduate students and is one of Georgia's four robust-tier institutions. It has been identified in each of the last five years as one of the Southeast's best public regional comprehensive universities.

The College of Education, under the direction of Dr. Kim K. Metcalf, offers baccalaureate, master's, specialist, and doctoral degrees across six departments: Counseling & Educational Psychology, Curriculum & Instruction, Educational Leadership & Professional Studies, Media & Instructional Technology, Physical Education & Recreation, and special Education & Speech language Pathology. The College is a recognized leader in the preparation of professionals in education, counseling, speech language pathology, and sport management. The College and constituent programs are nationally accredited, including NCATE, CACREP, and CAA.

The mission of the College of Education is to provide excellence in the initial and advanced preparation of professionals for a variety of settings; to foster an innovative, student-focused learning community; and to empower a faculty committed to teaching and the dissemination of knowledge. The College envisions itself as a student-focused educational community recognized for excellence in the comprehensive preparation of professionals who can positively impact school improvement.

To learn more about the University of West Georgia, visit our webpage at [www.westga.edu/coe](http://www.westga.edu/coe).

*Photograph courtesy of William Hogan.*

### *Call for Manuscripts*

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

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

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

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

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

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

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

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

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

**MWER Co-Editors**

Savilla Banister  
Mark Earley  
Julia Matuga  
Tim Murnen

*Bowling Green State University*

**MWER Editorial  
Advisory Board**

Kathleen Sullivan Brown  
*University of Missouri*

Tom Cody  
*Western Illinois University*

Carla Edlefsen  
*Ashland University*

Fran Falk-Ross  
*Northern Illinois University*

Sema Kalaian  
*Eastern Michigan University*

Mary Leahy  
*College of DuPage*

John Poster  
*University of Michigan-Dearborn*

Hema Ramanathan  
*University of West Georgia*

Sharon Valente  
*Ashland University*

Rachel Vannatta  
*Bowling Green State University*

Kathryn Wiggins  
*DePaul University*

Jane Williams  
*The Ohio State University*

Judy Zimmerman  
*Bowling Green State University*

**Layout and Design**

Judy E. Teska  
*JET Desktop Solutions*

***AYP Accountability Policy and Assessment Theory Conflicts***

*Randall S. Davies, Brigham Young University*

2

***Understanding School Board Members' Perceptions  
of Superintendents' Leader Behaviors***

*John V. Richard and Sharon D. Kruse,  
The University of Akron*

9

***An Exploratory Survey of Participants in Urban  
and Suburban Teacher Academy Programs***

*Leah Wasburn-Moses and Molly Kelly-Elliott,  
Miami University*

16

***The Effect of Supplemental Educational Services  
on Student Learning Outcomes***

*Jane Beese, University of Akron*

23

***The Impact of the Collective Efficacy of a School Community  
on Individual Professional Development Outcomes***

*Cynthia H. Geer, Xavier University  
Julie Q. Morrison, University of Cincinnati*

32

***Index of Articles: 2007–2008***

39

***Index of Authors: 2007–2008***

40

***Mid-West Educational Researcher Reviewers: 2007–2008***

*Inside back cover*

**MWER Publication Address**

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

**MWER Subscription &  
MWERA Membership Information**

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

---

# *AYP Accountability Policy and Assessment Theory Conflicts*

Randall S. Davies  
Brigham Young University

## ***Abstract***

*A major objective of NCLB is to hold schools accountable for student achievement including closing the achievement gap and raising standards of student academic proficiency. While the majority of individuals support the underlying values of NCLB, not all agree its accountability policy is reasonable. The most problematic issue is the mandate that schools be held accountable to ensure all students perform at grade level when the methods used by states to determine grade level proficiency preclude the attainment of this standard. The unintended consequences of a policy that does not align theoretically with established educational best practice inevitably result in frustration of educational practitioners and often unwarranted condemnation of the educational system.*

The Elementary and Secondary Education Act (ESEA) of 2002 (Public Law 107-110), commonly referred to as No Child Left Behind (NCLB), represents an extensive reform of educational expectations and requirements. Of primary importance, NCLB made a substantial change in the rules that govern schools by requiring that they describe their success primarily in terms of student performance. According to the U.S. Department of Education (2002), a major objective of NCLB is to hold schools accountable for closing the achievement gap between various groups of students, with the goal that all students obtain a proficient level of achievement on state standardized tests.

NCLB holds educators accountable through a policy that requires each school to make adequate yearly progress (AYP) toward the goal of improving student achievement for all groups of students in their care. To measure AYP, the legislation mandates that states administer high quality annual assessments to every child at specified grade levels. These assessments must be aligned to standards consistent with nationally recognized professional and technical standards; they must be administered in a valid and reliable manner, and they must test higher order thinking skills using multiple measures (U.S. Department of Education, 2002). All public schools are expected to meet NCLB goals, testing, and reporting requirements. Schools receiving federal Title I funds are additionally subject to a series of corrective actions should they fail to meet AYP.

ESEA legislation has been in effect for over five years now, and while the majority of individuals support the underlying values of NCLB, not all agree that the law's implementation has had beneficial outcomes (Bracey, 2006; Center on Education Policy, 2006; Fuller, 2006; McElroy, 2005; Mathis, 2003; Weaver, 2007). At issue is the unintended consequences of establishing educational policy that is not consistent with educational assessment theory and best practice; more specifically, the policy mandating that schools be held accountable for student's achievement, and the unattainable expectation that all students perform at grade level on standardized tests (Linn, 2004; Raudenbush, 2004; Welner, 2005). While these expectations may seem reasonable, assessment methods used by states to determine grade level proficiency preclude the

attainment of this standard. The unintended consequence of a policy that does not align theoretically with established educational best practice understandably has resulted in frustration for educational practitioners and often unwarranted condemnation of the educational system in general.

## **Determining Proficiency and Academic Achievement Standards**

While NCLB mandates that all students should obtain a proficient level of achievement on state standardized tests, and has targeted the year 2014 for reaching this goal, the act allows each state to determine its own performance standards and definition of proficiency (U.S. Department of Education, 2004). As might be expected, the definitions and levels of proficiency differ widely from state to state (Rosenburg, 2004).

Performance standards, or student academic achievement standards as they are referred to in NCLB (P.L. 107-110, Section 1111(b)(1)(A)), are expected to be challenging; however, they must also be reasonable (Airasian & Russell, 2008; Linn, 2004). A requisite starting point for the process of establishing performance standards is a credible analysis of typical student performance in order to determine what can reasonably be expected from students at each grade level. Setting meaningful performance standards requires the use of established grade level norms, experience, and good judgment. It assumes that instruction at each grade level is based on a standardized curriculum that is aligned with state academic content standards. Establishing developmentally and educationally appropriate levels of performance is not always easy. Setting student achievement expectations too high or too low is undesirable in terms of motivation and obtaining increases in performance (Torrance, 1970). This is true for both teachers and students. There are political consequences to consider as well. Setting the standards too high will result in large numbers of students failing to pass state assessments. Setting standards too low invites criticism that students graduate without having to learn anything. Both situations are undesirable. Not only is setting performance standards a complicated process, it is clear that intelligent

---

people disagree on what the definition of proficient should be (Linn, 2003; Rosenburg, 2004).

State standardized tests are carefully created to assess student learning of content standards. This is important if the tests are to be valid estimates of student ability. Content standards tell teachers what will be tested and what they should teach. However, content standards do not indicate how well a student must perform in order to be considered proficient; performance standards are needed to make determinations regarding proficiency (Linn & Miller, 2005).

A common procedure for determining what a proficient grade level performance, or performance standard, should be is to have a group of educational experts and psychometricians make this decision (Airasian & Russell, 2008; Winchester, 2006). Based on their experience, these experts examine grade level norms, evaluate what students should be able to do at that age, consider appropriate expectations, and then set cut points for passing each exam. Cut points (i.e., passing scores) are typically reexamined on a regular basis to ensure they are appropriate, reasonable, and fair (Winchester, 2006). This method of determining standards for student performance adheres to accepted assessment theory and practices (Airasian & Russell, 2008; Linn & Miller, 2005; Reynolds, Livingston & Willson, 2006). What this means is that not everyone who takes the test will pass; it is an impossibility when the passing score is based on an analysis of typical performance.

### *Criterion and Norm Referenced Assessments*

A common error in understanding for many people suggests that because state standardized tests are supposed to be criterion referenced, not norm referenced, the use of a cut point to indicate proficiency does not affect the number of students who can achieve proficiency. If standardized tests were designed to test mastery, this may be true; but in practice, this assumption is incorrect on several accounts. Most importantly, the concept of criterion reference refers to how the results are interpreted and reported, not how a proficiency cut point is determined, and thus, how many students would reasonably be expected to pass the test (Airasian & Russell, 2008; Linn & Miller, 2005; Reynolds, Livingston & Willson, 2006).

State standardized tests are considered to be criterion referenced due to the fact that individual results are interpreted against content standards rather than the performance of other students taking the exam each year. Unfortunately, state tests are typically not designed to maximize this type of interpretation. State standardized tests typically cover a large domain of learning tasks across multiple standards which limits the number of questions that can be asked. Test items are selected based on their discriminating power and typically exclude items that are too easy or too difficult (Winchester, 2006). This type of test design is better suited to norm referenced interpretations (Linn & Miller, 2005); however, state tests are designed in this way to accommodate item response theory requirements (Indiana Department of

Education, 2007). In order to make use of item response theory, test makers routinely and justifiably use typical performance analysis and grade level norms to determine item difficulty and discriminating power rather than using absolute measures of content mastery. As a result, even though the interpretation and reporting of these test results are criterion referenced, acceptable individual student performance is determined by a comparison of typical student performance based on grade level norms expectations.

This is not a criticism of state examinations. Based on the evidence, these tests seem to provide valid and reliable estimates of student ability. At issue is how the results are used (i.e., consequential validity); more specifically, while one might hope that all students in a particular grade would be able to master the curricular content, because passing scores are based on an analysis of typical student performance, it is unreasonable to expect all students would pass state standardized tests. Given the way passing scores are determined, some students will always be below grade level on state standardized tests, and the number of students below grade level will be fairly constant. Because AYP is based on the percentage of students passing state standardized tests—and it is impossible for all students to pass the exam given the method for determining passing scores—inevitably all schools will eventually fail to meet increasingly more challenging AYP benchmarks.

### *Expected Results When Using Typical Performance to Determine Proficiency*

In theory, the practice of establishing appropriate grade level performance standards is based on the belief that some human characteristics (e.g., intelligence, cognitive ability, and academic achievement) are normally distributed in the population. There will always be a few individuals who are considerably more able or skilled than the average; and a few individuals who will always be significantly less able or skilled compared to the average; however, a large percentage of students (68%) will demonstrate average performance (i.e., one standard deviation above or below the mean). Statistically, students who score in this range are arguably quite similar given the fact that all tests are prone to a certain amount of measurement error and an individual's true scores must be considered within a derived confidence interval (Linn & Miller, 2005). Still, a cut point must be established, and inevitably it will exclude some students from passing the exam.

Students who pass state tests are deemed proficient; students who fall significantly below the range of average performance established for a specific assessment are by definition considered to be below grade level. The concept of grade level proficiency suggests a band of acceptable performance based on the observed typical performance of students in that grade (American Federation of Teacher, 2004; Linn & Miller, 2005). If proficiency is defined as those students who obtain a passing score on a test, and the passing score

is determined through an analysis of typical performance for students in that grade, then it is only reasonable to expect that a group of students will always be excluded from being labeled proficient. The exact percentage of students who fail to reach a proficient level of achievement will depend on where the cut point is set.

Obviously, if the passing score of a test is set to equal the mean or average score obtained by students taking the test, by definition half of the students (50%) would pass and the other half would fail. It would be unreasonable to expect that all students pass the test if this were the case, because by definition it is impossible for more than 50% of the students taking any given test to be above average. However, a very similar thing happens when a cut point is used to determine proficiency instead of the mean.

An illustration of this phenomenon is presented in Table 1. Since all states use different assessments and set different cut points, it is more appropriate to look at each state individually. For the purposes of this paper, the state of Indiana is used as a case study. Table 1 data represent pass rates for the third grade Language Arts and Mathematics assessment in the state of Indiana over the past five years; similar performance patterns can be observed in other states and for other grade levels.

Table 1  
*Statewide 3<sup>rd</sup> Grade Language Arts Assessment Results Comparison by Year*

Year	Mean	SD	Z score of Passing Score	% Passing
2002	436.7	61.8	-0.53	72
2003	442.4	65.5	-0.59	74
2004	443.0	63.5	-0.61	75
2005	442.6	64.9	-0.59	75
2006	438.9	67.3	-0.56	74
2007	442.1	64.2	-0.60	75
Average	441		-0.58	74

*Statewide 3<sup>rd</sup> Grade Mathematics Assessment Results Comparison by Year*

Year	Mean	SD	Z score of Passing Score	% Passing
2002	416.0	61.2	-0.38	67
2003	424.0	63.7	-0.49	71
2004	426.5	62.7	-0.53	73
2005	426.1	64.8	-0.51	73
2006	425.0	65.8	-0.49	72
2007	420.5	64.2	-0.43	70
Average	423		-0.47	71

Established cut points for each year was 393 for Math and 404 for Language Arts

As assessment theory predicts, in both situations you get a consistent percentage of students passing the exam each year. Variations in the result are just as likely explained by differences in the exams given each year, teachers better aligning their instruction with state content standards, or by

schools better preparing students to take exams as they are by any real change in the ability of the students from year to year. Analysis of scores from the National Assessment of Educational Progress (NAEP) confirm this result suggesting no significant overall improvement in student achievement has occurred since the enactment of NCLB legislation (Bracey, 2006; Center on Education Policy, 2006). Given that in theory achievement and scholastic ability are believed to be normally distributed and that the criteria for establishing specific proficiency cut points are based on an analysis of typical performance, this result would be expected (Armor, 2006; Reynolds, Livingston & Willson, 2006). If anything, this situation suggests that the assessments used are somewhat reliable at measuring what they measure.

In general, achievement trends in each state show the overall percentage of students that pass state standardized assessments each year has been, and will likely remain, fairly constant for students as an overall group. One thing is clear; equating school quality with the percentage of students at that school who achieve “proficiency” does not withstand serious scientific scrutiny when the standard for proficiency is based on an expectation of typical grade level performance (Raudenbush, 2004). The assumption that schools can somehow cause all students to perform at or above grade level proficiency constitutes an unrealistic expectation (Linn, 2004; Welner, 2005).

## Breakdown of Results

There are many reasons why children are left behind in our schools; lacking the opportunity to receive a high-quality education is but one (Reigeluth & Beatty, 2003).

A well-substantiated body of research links teacher quality and student learning (Darling-Hammond, 2000; Mendro, 1998; Stedman, 1997; Wenglinsky, 2002), with a particularly strong positive correlation between teacher subject area certification and student achievement (Goldhaber & Brewer, 1996). However, being highly qualified and providing a quality learning opportunity does not mean teachers will be completely effective for all students in all situations. Meeting AYP targets is not a valid indicator of teacher and school quality; there are better explanations for understanding poor student performance in schools (Raudenbush, 2004).

Continuing with the Indiana example, consider a comparison of three school districts. In Indiana, school districts are called school corporations. Table 2 shows a demographic comparison for three school corporations all located in one metropolitan area which includes the surrounding rural population. These school corporations are largely divided along lines of Socio-Economic Status (SES) with larger minority populations and a higher proportion of special needs students in the low SES areas. School Corporation A is low achieving, Corporation B has moderate achievement, and Corporation C is high achieving (see Table 3). School Corporation A failed to meet their AYP targets in the 2006-2007 school year. Both the other school corporations met AYP goals that year. These

three school corporations are quite diverse, yet together they constitute a reasonably representative sample of the state. Although the corporations combined have slightly more minorities, more students receiving free or reduced lunch, and more students with special needs, they correspondingly have a slightly lower percentage of students passing the state assessments in language arts and mathematics. While there is some variation in the overall result, as expected, the average number of students in the state passing the test and the average number of students in these combined school corporations passing the test has varied little over the past five years (see Figures 1 and 2).

All teachers in the state of Indiana are required by law to be highly qualified to teach in their subject areas, yet these three corporations have consistently obtained very different results. On the surface, based on the percentage of students passing state tests, School Corporations A and B might be seen as less effective. Some might erroneously believe that the teachers in these schools are less qualified, dedicated, or able. However, looking at the trend analysis in Figures 1 and 2, these two school corporations had the greatest gains in achievement over the past five years.

One important aspect related to school success is the portion of special needs students being served by a school. Schools with large proportions of special needs students are typically at greatest risk of failing to meet AYP expectations (Kim & Sunderman, 2005). In Indiana during the 2006-2007 school year,

44 (71%) of the 62 school corporations that failed to meet AYP targets failed in the special needs category. Thirty (48%) of the 62 corporations failed only in this area. This is not surprising given that special need students understandably do less well on tests than their more able peers. When compared to regular students, special needs students will always end up at the low end of the assessment distribution.

Another pattern in these data is the correlation between the number of students receiving Free or Reduced Lunch and academic achievement. Again, schools with large proportions of students living in poverty are at greater risk of failing to meet AYP expectations (Marzano, Pickering, & Pollock, 2005; Holmes-Smith, 2006; Heck, 2006). Notably, minority students are more likely to be living in poverty; however, this result might suggest that it is not a students' minority status that influences scholastic performance, but rather the fact that they live in poverty and are more likely to have special learning needs. Teaching students under these circumstances represents a special challenge for schools (Heck, 2006).

Given this pattern of achievement, none of these school corporations are on track to have all their students at a proficient level of achievement on state standardized tests by the 2013-2014 school year. The only logical conclusion one can make is that each of these school corporations will eventually and inevitably be labeled as "failing" (Armor, 2006). Regardless of the reasons students fail to meet grade level expectations, eventually all schools will fail to meet the ever

Table 2  
*School Corporation and State Demographics for 2006-2007*

	Total Enrollment	% Minorities	% Free/Reduced Lunch	% Special Needs	Graduation Rate	Drop Out Rate
Corporation A	21874	68	62	24.6	60.4	22.8
Corporation B	5682	22	48	20.4	52.6	18.6
Corporation C	10612	13	15	14.6	83.2	15.0
Combined	38168	46	47	21.2	65.6	20.0
State	1045702	33	38	17.8	76.5	11.2

Table 3  
*Percent Passing Language Arts Assessment by School Corporation and Year*

	2002	2003	2004	2005	2006
Corporation A	51.3	50.8	55.0	54.3	56.8
Corporation B	63.6	62.4	65.9	67.1	69.6
Corporation C	81.6	82.7	82.7	83.1	84.4
Combined	61.2	61.4	64.2	64.1	66.4
State	59.4	70.6	71.2	71.7	71.4

*Percent Passing Mathematics Assessment by School Corporation and Year*

	2002	2003	2004	2005	2006
Corporation A	51.8	53.0	56.4	57.5	58.5
Corporation B	68.3	70.3	72.3	75.1	75.5
Corporation C	83.7	85.3	85.3	86.2	85.7
Combined	62.8	64.6	66.7	68.0	68.6
State	68.0	71.4	72.2	74.0	73.7

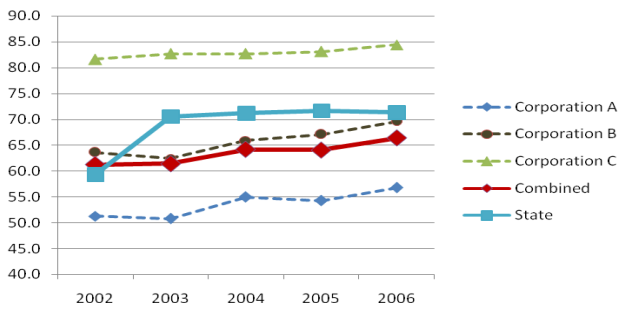


Figure 1. Percent Passing Language Arts Assessment by School Corporation and Year

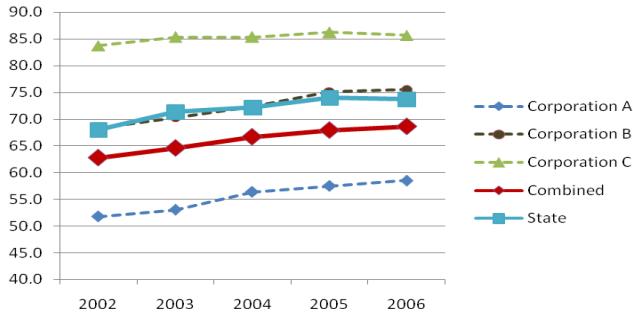


Figure 2. Percent Passing Mathematics Assessment by School Corporation and Year

increasing targets of AYP. Because of the way proficiency is defined and measured, even schools recognized for excellence will eventually succumb.

### Inevitable and Unanticipated Consequence of Impossible Expectations

Making adequate yearly progress (AYP) is primarily defined as the degree to which schools meet specific targets in the number of students at or above proficiency on state examinations (American Federation of Teacher, 2004). This applies to all students at the school, including disaggregated groups of students within a school. In order to attain a positive AYP status and to avoid the “failing school” label, many schools have at times resorted to exceptional practices in an attempt to maximize student achievement on NCLB mandated standardized tests (Orina & Davies, 2006). Teachers often note feeling relieved once standardized testing is done each year so they could get back to the normal tasks of educating children. One unintended consequence that is becoming harder for teachers to overcome is the stress, frustration, and anger they feel at being held accountable for things over which they have little or no control (Orina & Davies, 2006; Sulok, 2005).

Regardless of assurances that teachers would not be held individually responsible for a school’s failure to achieve AYP, many teachers feel they are being blamed for low test results in spite of their best efforts. Parents are understandably concerned when a school fails to meet AYP. The National Center for Educational Statistics (2007) estimates that approximately

27% of schools failed to meet AYP in the 2004-2005 school year nationwide. It is expected that an increasing number of schools will fail to meet AYP targets each year (Wiley, Mathis, & Garcia, 2005). As this happens, teacher morale can be expected to decrease even more.

This is especially true for teachers who work with students in various special needs categories. Often these students are held to the same performance standards as general education students even though they will eventually end up at the low end of the achievement distribution. It is clearly not reasonable to blame special needs students for this situation, nor is it reasonable to blame schools when they cannot reach unrealistic expectations of performance.

In many schools the special needs populations are increasing (Gunter, 2005). Schools with large populations of students in these categories often experience an enormous drain on resources attempting to provide for these students and ensure they achieve. In fact, students who do not perform adequately often become the main focus of instructional efforts to the exclusion of those groups of students already attaining proficient levels of achievement (Orina & Davies, 2006). Teachers in schools with large numbers of students identified in ENL or special needs categories often indicate a sense of futility at meeting the testing requirements (Orina & Davies, 2006). They are becoming overwhelmed with the challenge of making sure these students test well in comparison to the general education student population. Critics of NCLB argue that federal efforts demanding results-based accountability are presumptively futile because they assume that all students can and will perform at a proficient level academically regardless of their abilities and motivation (Schrag, 2004). One might argue that teacher morale is declining as a direct result of the pressure being put on them to meet unrealistic accountability standards that are becoming more and more unattainable (Sulok, 2005). In many ways the mandate to ensure all students achieve proficiency is a “shoot for the moon; even if you miss, you’ll land among the stars” philosophy. The fact that teachers and schools are being asked to meet unrealistic AYP goals does not seem to matter.

### Discussion Summary

In general, the methods for determining grade level proficiency used by most states are appropriate and in accordance with accepted assessment theory. Setting meaningful performance standards based on grade level proficiency requires the use of established grade level norms and good judgment. Certainly there is disagreement between states on what it means to set challenging student academic achievement standards, but this is to be expected given the nature of human beings and the often conflicting and varied views regarding the purposes of public education.

The most problematic aspect of NCLB policy is that the way states define grade level proficiency will always exclude some students from being categorized as proficient. This would not be a problem except for the fact that by equating



---

the cut point (i.e., passing score) on state standardized test with performance standards, and then expecting all students to be able to meet that standard, constitutes an unrealistic expectation and impossible burden for the educational system. Certainly some schools currently enjoy a greater level of success in terms of the number of students passing state assessments compared to other schools; however, all schools, even the best schools, will eventually reach an achievement barrier in which they cannot do any better in terms of meeting ever increasing AYP targets.

While most teachers feel it is acceptable to expect that their students attain an adequate level of achievement, there is a growing feeling of frustration among educators. Often this frustration seems to stem from the fact that teachers are being held accountable for factors and conditions they have no ability to change. Important factors that affect learning include: regular attendance at school; the effort and attention students put into their studies; the support and encouragement students receive from home; the economic and societal influences they experience; and the students' abilities, willingness, interests, and intentions for the learning expected of them as students. These and other issues are seen by teachers as factors they cannot control. Unfortunately, AYP status is considered by some to be synonymous with school quality. State assessments of student ability are not valid indicators of school and teacher quality. Despite the best efforts of excellent teachers, many students do not attain proficient levels of achievement, and even in the best classrooms, not all students achieve excellence.

### Conclusions

Assessment is a fundamental aspect of the teaching and learning process; however, a proper understanding of assessment is essential when establishing educational policy. Policy that does not align theoretically with established educational best practice inevitably results in frustration of educational practitioners and often unwarranted condemnation of the educational system. One aspect of the current ESEA's accountability policy that does not align with educational theory is the expectation that all students can meet grade level standards when such standards of proficiency are based on the normal distribution of typical student performance for that grade level. Federal mandates that expect schools to meet arbitrary AYP targets should be removed from NCLB accountability policy. Students should be tested, but it is inappropriate to use results of state standardized assessments as the primary evidence for judging the quality of teachers and schools.

NCLB seems to operate on the premise that every child can learn and will learn if they are provided with a highly qualified teacher and a beneficial learning environment that utilizes scientifically proven practices. However, there are many reasons why children are left behind in our schools; lacking the opportunity to receive a high-quality education is but one (Reigeluth & Beatty, 2003). Educational accountability policy needs to hold teachers and schools accountable

for things they have control over. It is the responsibility of educators to provide quality instruction within a beneficial learning environment. They should care about their students and do what they can, within reason, to help them learn. It is the students' responsibility to take advantage of the opportunities they are given. Students should be held accountable for their own achievement. Assessment results help teachers and students identify areas for improvement. While there are likely several things schools and individual teachers can do to improve instructional practices, meeting unrealistic targets for student achievement when those targets are by definition impossible to reach is something schools should not be burdened with.

### References

- American Federation of Teacher (2004). *Eight misconceptions about the No Child Left Behind act's (NCLB) adequate yearly progress (AYP) provisions*. Retrieved November 30, 2006, from <http://www.aft.org/topics/nclb/downloads/8Misconceptions.pdf>
- Airasian, P. W., & Russell, M. K. (2008). *Classroom assessment: Concepts and applications* (6<sup>th</sup> ed.). Boston: McGraw-Hill.
- Armor, D. J. (2006, August). Can NCLB close the achievement gap? *Teachers College Record* (ID Number: 12667). Retrieved August 23, 2006, <http://www.tcrecord.edu.org>
- Bracey, G. W. (2006). The 16<sup>th</sup> Bracey report on the condition of public education. *Phi Delta Kappan*, 88(2), 151-166.
- Center on Education Policy (2006). *From the capitol to the classroom: Year 4 of the No Child Behind Act, summary and recommendations*. Retrieved November 23, 2006, from <http://www.cep-dc.org/nclb/Year4/NCLB-Year4Summary.pdf>
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives*, 8(1). Retrieved March 30, 2007, from <http://epaa.asu.edu/epaa/v8n1>
- Fuller, B. (2006). *Is the No Child Left Behind act working?* Berkeley, CA: Policy Analysis for California Education.
- Goldhaber, D. D., & Brewer, D. J. (1996). Evaluating the effect of teacher degree level on educational performance. *ERIC Digest*. (ERIC Document Reproduction Service No. ED406400).
- Gunter, C. (2005). *Educating students with limited English proficiency*. Unpublished masters project. Indiana University, South Bend.
- Heck, R. H. (2006). Assessing school achievement progress: Comparing alternative approaches. *Educational Administration Quarterly*, 42, 667-698.
- Holmes-Smith, P. (2006). *School socio-economic density and its effect on school performance*. Retrieved October, 21, 2007, from [http://www.mceetya.edu.au/verve/\\_resources/SES\\_Report.pdf](http://www.mceetya.edu.au/verve/_resources/SES_Report.pdf)
- Indiana Department of Education. (2007). *Guide to test interpretation*. Retrieved October 30, 2007, from [http://www.doe.state.in.us/istep/pdf/GTI/49640-W\\_GTI\\_s07IN.pdf](http://www.doe.state.in.us/istep/pdf/GTI/49640-W_GTI_s07IN.pdf)

- Kim, J. S., & Sunderman, G. L. (2005). Measuring academic proficiency under the No Child Left Behind act: Implications for educational equity. *Educational Researcher*, 34(8), 3-13.
- Linn, R. L. (2003). Accountability: Responsibility and reasonable expectations. *Educational Research*, 31(7), 3-13.
- Linn, R. L. (2004). *Rethinking the No Child Left Behind accountability system*. Paper presented at a forum sponsored by the Center on Education Policy, Washington, DC.
- Linn, R. L., & Miller, M. D. (2005). *Measurement and assessment in teaching* (9<sup>th</sup> ed.). Upper Saddle River, NJ: Pearson Education.
- Mathis, W. J. (2003). *No Child Left Behind: Costs and benefits*. Phi Delta Kappa. Retrieved on August 7, 2006, from <http://www.pdkintl.org/kappan/k0305mat.htm>
- Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2005). *Classroom instruction that works*. Upper Saddle River, NJ: Pearson Education.
- McElroy, E. J. (2005, January). *Teaching left behind. Teaching Pre K-8*. Retrieved August 7, 2006, from [http://www.teachingk8.com/archives/teacher\\_to\\_teacher/teaching\\_left\\_behind\\_by\\_edward\\_j\\_mcelroy.html](http://www.teachingk8.com/archives/teacher_to_teacher/teaching_left_behind_by_edward_j_mcelroy.html)
- Mendro, R. L. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) database: Implications for educational evaluation and research. *Journal of Personal Evaluation in Education*, 12, 247-256.
- National Center for Educational Statistics (2007). State educational reforms: Data tables. Retrieved October 10, 2008, from [http://nces.ed.gov/programs/statereform/saa\\_tab6.asp](http://nces.ed.gov/programs/statereform/saa_tab6.asp)
- No Child Left Behind Act of 2001, Pub. Law No. 107-110, 115 Stat. 1425 (2002).
- Orina, E., & Davies, R. (2006, Mar). *The unintended consequences of NCLB*. Presented at the annual conference of the Michigan Academy of Science, Arts, and Letters (MASAL), Rochester, MI.
- Raudenbush, S. W. (2004). Schooling, statistics, and poverty: Can we measure school improvement? *Educational Testing Service, William H. Angoff Memorial Lecture Series*. 6.
- Reigeluth, C., & Beatty, B. (2003). Why children are left behind and what we can do about it. *Educational Technology*, 43(5), 24-32.
- Reynolds, C. R., Livingston, R. B., & Willson, V. (2006). *Measurement and Assessment in Education*. Boston: Allyn and Bacon.
- Rosenburg, B. (2004). *What's proficient? The No Child Left Behind act and the many meanings of proficiency*. Retrieved November 30, 2006, from <http://aft.org/pubs-reports/downloads/teachers/WhatsProficient.pdf>
- Stedman, L. C. (1997). International achievement differences: An assessment of a new perspective. *Educational Researcher*, 26, 4-15.
- Shrag, P. (2004). *Bush's education fraud. The American prospect*. Retrieved on August 7, 2006 from <http://www.prospect.org/web/page.wv?section=root&name=ViewPrint&articleId=6998>
- Sulok, N. (2005, December 12). MEA President says "No Child" hurting morale. *South Bend Tribune*. Retrieved on November 30, 2006, from <http://www.southbendtribune.com/apps/pbcs.dll/article?AID=/20051212/News01/512120316/0/BIZ>
- Torrance, E. P. (1970). *Encouraging creativity in the classroom*. Dubuque, IA: Wm. C. Brown.
- U.S. Department of Education, Office of Elementary and Secondary Education (2002). *No Child Left Behind: A desktop reference*, Washington, DC.
- U.S. Department of Education (2004). *A guide to education and No Child Left Behind*. (Office of the Secretary, Office of Public Affairs). Washington, DC.
- Weaver, R. (2007). NCLB: The view up close. *Social Policy*, 37(3/4), 45-48.
- Welner, K. G. (2005). Can irrational become unconstitutional? NCLB's 100% presuppositions. *Equity and Excellence in Education*, 38, 171-179.
- Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives*, 10(12). Retrieved June 20, 2008, from <http://epaa.asu.edu/epaa/v10n12/>
- Winchester, D. (2006, October). *Indiana academic standards and benchmarks*. Educational technology meeting presentation, Indianapolis, IN.
- Wiley, E., Mathis, W., & Garcia, D. (2005). *The impact of the adequate yearly progress requirement of the federal "No Child Left Behind" act on schools in the Great Lakes region*. Retrieved November 30, 2006, from [http://greatlakescenter.org/g\\_l\\_new\\_doc/EPSE-0505-109-EPRU\\_EXEC.pdf](http://greatlakescenter.org/g_l_new_doc/EPSE-0505-109-EPRU_EXEC.pdf)

---

# *Understanding School Board Members' Perceptions of Superintendents' Leader Behaviors*

John V. Richard  
Sharon D. Kruse  
The University of Akron

## *Abstract*

*The study of school leadership has been a topic of considerable investigation. Primarily, this literature has focused on the leadership of principals and superintendents. Although school boards work hand in hand with school leaders concerning the decision making functions of district business, a literature base specifically focused on school board members and their perceptions of school superintendents' leadership behavior is lacking. This paper provides a framework for understanding school board members' perceptions and suggests that school board members' views of superintendents' leadership behaviors may be influenced by demographic factors including years of experience and gender of board members.*

## Introduction

The relationship between a public school superintendent and his or her school board is critical. The board of education is responsible for hiring, evaluating, and compensating the superintendent. In turn, the superintendent is required to keep the board apprised of important operational and instructional issues within the district. Ideally, this relation is harmonious; board members and the superintendent work together making decisions, setting vision and mission, and solving problems, all with an eye toward district improvement and student learning. Unfortunately, this outcome is not common nor is it guaranteed.

Typically, central office staff is hired from the ranks of building leaders—elementary, middle and high school principals. From there, the path to the superintendent's office is most often through the district office; former assistant superintendents, curriculum directors and business managers all “move up” to take on the role. Certainly, these roles provide some of the necessary background knowledge and skills for success in a school district's highest office. However, personnel in these positions are often buffered from direct contact with school board members. It is often not until someone is hired into the position of superintendent that they are required to have direct and on-going contact with school board members (Houston & Eadie, 2005).

Recent research suggests that because of this buffering, numerous superintendents have never acquired the skills that will ensure a strong superintendent-board relationship (Goodman & Zimmerman, 2000; Houston & Eadie, 2005). The result is often a dissatisfied, frustrated, and/or angry board (Houston & Eadie, 2005). These administrators, according to Houston and Eadie, had performed well in other administrative roles in the district, but these roles did not prepare them for the consistent interaction with board members required

by the position of superintendent. Of course, when the board chooses to hire these leaders they had full confidence in their abilities. Yet, as time and work progress, these initially rosy perceptions often sour, leaving board members disappointed and unsatisfied with the performance of the superintendent. Once displeased, it is often difficult for the board to remain confident in the school leader's abilities and effectiveness. In this way, board of education members' perceptions of their superintendent's leadership behavior are important and have the potential to affect a superintendent's efficacy.

Foundational to the discussion of school board and superintendent relations is the understanding of the ways in which school board members perceive a superintendent's behaviors and actions. Captured in the literature under the more general rubric of communication or public relations, school board members' perceptions of superintendent leadership behaviors have been under-represented in the literature of school leadership. The paper seeks not only to develop ideas related to school board perceptions of superintendent leadership actions and behaviors, but seeks to do so by comparing observed behaviors with those considered ideal.

Two general research questions guided the direction of the study. Are there significant differences in the perceptions of school board members regarding ideal leader behaviors (“what should be”) and the actual (real) leader behaviors (“what is”) of their school superintendents? Secondly, what are the factors that contribute to the differences in these perceptions; are these factors associated with demographics? To answer these questions, this study focuses on the following research objectives:

1. The development of understanding related to school board members perceptions of actual (real) and ideal superintendent behaviors, and;
2. The utility of the findings to aid in understanding organizational outcomes and performances.

---

## Leadership in practice

Leadership as a social and educational phenomenon has been the subject of considerable attention in the literature of business, sociology, psychology, and education. Common to these studies is an attempt to define and type leader behaviors, characteristics, and actions (Bass, 1990; Yukl, 2002). There may be some difference in the specifics of a definition, but most authors conclude that leadership generally comprises *the exercise of intentional social influence through which members of a group are steered toward a goal through a process of structured activities, efforts, and individual or shared endeavors* (Bass, 1990; Bryman, 1986; Yukl, 2002). Foundational to thinking about leadership has been a concern for the production-focused work tasks in which leaders engage, as well as the people with whom they engage to perform these tasks.

The tandem theoretical constructs of consideration (people-related behavior) and initiation of structure (task-related behavior) have served to delineate the landscape of thinking about the work of leaders since the last century (Bass, 1999; Conger, 2004; Weick, 2001). Early research suggested that consideration addressed the social and emotional needs of organizational members, including recognition for their efforts, satisfaction with the work environment and task as well as other less tangible aspects of organizational culture and climate (Fleishman, 1973, 1995; Yukl, 2002). Research into those aspects of leadership thought to be initiating structures included leadership activities such as strategic planning and organizing, definition of work tasks and products, and evaluating individual and organizational progress toward goals (Fleishman, 1973, 1995; Tollerico, 1989; Yukl, 2002). The attention to these constructs as separate measures produced little in the way of concrete understandings concerning effective leadership, but when considered as associative and mutually informing notions they proved useful in the practical delineation of the disparate activities and actions of leaders. By considering these dual orientations of leadership behaviors, both employers and researchers began to consider the ways in which these two constructs interacted to create ideal organizational outcomes.

In considering the construct of initiating structures, research suggested the significance of the task as well as a leader's identification with the core functions of school improvement and progress (Gronn, 2003; Ogawa & Bossert, 1995). When leaders focused their attentions on actions related to developing and implementing a vision, creating and adopting policy, practices and procedures for the day to day work of school and district personnel and monitoring and evaluating progress toward organizational goals it was thought that they were successfully contributing to the school and district goal attainment (Hargreaves & Fink, 2006;

Rowan, 2002; Spillane, 2006). However, research suggests that leaders, who only attend to the tasks of leadership, ignoring those with whom they work and rely, were less successful than leaders who thought about their charges in a more comprehensive manner (Kotter, 1995). In turn, those behaviors related to the well-being of the people who work within the district such as providing encouragement and recognition, communication of meaningful information in a timely and clear manner as well as openness and consultative behaviors, were considered critical for understanding leadership practices in schools and districts (Louis, 1994; Tollerico, 1989; Vechio, 2006).

What scant research exists concerning the topic of board members' perceptions of school superintendents' leader behaviors suggests that a positive board-superintendent relationship, including the board's ability to maintain a positive perception of the superintendent, is critical to the superintendent's effectiveness (Goodman & Zimmerman, 2000; Harrington-Lueker, 2002; Hoyle & Skrla, 1999; Peterson & Short, 2001). This importance appears rational enough, given the fact that the board of education has authority to hire, fire, reward, renew contracts, and reinforce the work of the superintendent. Kowalski (1999) asserts that rapid turnover in the superintendency is often attributed to poor relationships between a superintendent and school board members. Dillon and Halliwell (1991) found that when superintendents' perceptions of his or her own purposes, strengths, and weaknesses were similar to those of board members, superintendents were more likely to be retained regardless of other performance data. Congruence in the perceptions is emphasized as a critical factor for ongoing superintendent effectiveness.

However, the literature on school board/superintendent relationships is lacking in empirical studies of the phenomena that contribute to the development of board perceptions of effectiveness. As has been discussed above, much of the prior work (Goodman & Zimmerman, 2000; Harrington-Lueker, 2002; Hoyle & Skrla, 1999; Houston & Eadie, 2005; Peterson & Short, 2001) has focused on documenting that perception matters and that superintendents should take into account the perceptions of board members when considering their actions. It is important that research be completed that tease out the nuances that inform and form the perception forming process.

## Methods

To address the gap in the literature, it was decided that an empirical study addressing the perceptions of superintendent effectiveness would be completed. Survey instruments were chosen for this study based on the study's focus of superintendents' leader behaviors from board members' perspectives. Prior work in this area has employed the *Leader Behavior*

---

*Description Questionnaire* (LBDQ) to measure two main constructs, Initiation of Structure and Consideration (Fleishman, 1995). In this research, consideration was defined as the degree to which a leader shows concern and respect for followers, looks out for their welfare, and expresses appreciation and support (Bass, 1990). In short, Consideration represents the people-skills of leader behavior. Initiating Structure, or Initiation of Structure, is the degree to which a leader defines and organizes the leader's personal role and the roles of followers, is oriented toward goal attainment, and establishes well-defined patterns and channels of communication (Fleishman, 1973, 1995). Initiation of Structure represents the production or task behavior of leadership. Consideration and Initiating Structure have been considered to be among the most robust of leadership concepts (Fleishman, 1995).

In keeping with the prior research on leadership behavior (Fleishman, 1995), this study used the Leader Behavior Description Questionnaire-Form XII (LBDQ-Real), which measures perceptions of the actual behaviors observed in superintendent leadership behavior. The Ideal Leader Behavior Description Questionnaire-Form XII (LBDQ-Ideal), which measures those behaviors a board member would perceive as ideal leadership actions, and a Personal Data Sheet were also used for the collection of data in this study. The LBDQ-Form XII was developed by Stogdill (1963) and continually updated to account for changes of job role and vocabulary over time (Fleishman, 1995; Judge, Piccolo, & Ilies, 2004). The LBDQ—Form XII instruments set forth items that measure the perceptions of a leader's ability to attend to people within the district environment (known as the Consideration sub-scales) and his or her ability to attend to tasks within the district (known as the Initiating Structure sub-scales).

The two subscales of Consideration and Initiating Structure have been widely used in empirical research. Consideration and Initiating Structure have been considered to be among the most robust of leadership concepts (Fleishman, 1995). A meta-analysis of the relationship of Consideration and Initiating Structure with leadership provided support for the validity of these two subscale constructs in conducting further leadership research (Judge, Piccolo, & Ilies, 2004), revealing that both Consideration and Initiating Structure have main effects on numerous criteria noted in the leadership literature as fundamental indicators of effective leadership. Reliability of the LBDQ-Form XII appears relatively strong. Internal consistency coefficients were reported between .70 and .80, using a modified Kuder-Richardson formula (Stogdill, 1963).

The study specifically focused on the effects that the independent variables of gender, educational level, and years of experience of board members have on the dependent variables of board members' perceptions of superintendents'

leader behavior as measured on the LBDQ-Form XII subscales of Initiation of Structure and Consideration, both real and ideal. Additionally, the effects of the independent variables of gender and years of experience of superintendents on board members' perceptions of superintendents' leader behavior (dependent variable) were analyzed.

School board members from public school districts throughout a mid-western state were selected for the study. The following sampling procedures were utilized. School districts were divided into three categories: student enrollments below 2,000; student enrollments between 2,000 and 4,000; and student enrollments greater than 4,000. Following this non-proportional stratification, 50 districts from each category were randomly selected using the "sampling with replacement" method. Using this method permitted every district within each stratified population to be given an equal chance of being selected and therefore every possible sample within each category was equally probable. As the name of each school district was drawn from the population for the sample, the name of the district was then recorded and subsequently returned, guaranteeing that each school district had an equal chance for selection to the study. Individual school board members from each of the selected districts were provided an opportunity to participate in the study.

Data in this study were collected from school board members in public school districts, utilizing survey research specifically with self-administered questionnaires. The school board president from each of the selected school districts received by regular mail an initial letter of invitation and explanation, followed two weeks later with a reminder via email to the board president and superintendents encouraging participation, and then a packet of materials was mailed to each board president one week following the email reminder. One month following the mailing of packets, an email was sent to superintendents in all districts in which fewer than two responses had been completed and returned. Two weeks following this email, phone calls were made to districts in which there were no responses, for the same purpose of encouraging board members' participation in the study. The projected sample size was minimally 750 (3 categories by size, multiplied by 50 selected districts in each category, multiplied by at least 5 board members per school district). One hundred ninety-nine school board members completed and returned a survey for purposes of the research study that translated into a 27% response rate. Tables 1-3 provide demographic information of board members, superintendents, and school districts, respectively.

## Findings

Two-way multivariate analysis of variance (MANOVA) was utilized in the hypotheses that compare the mean scores

Table 1  
*Demographic Information of Participants Analyzed*

	Demographic Variable	Frequency	Percentage
Total	All respondents	199	100
Gender	Male	115	57.8
	Female	84	42.2
Service	≤ 2 years experience	56	28.1
	3-5 years experience	45	22.6
	6-9 years experience	53	26.6
	≥ 10 years experience	45	22.6
Ed. level	H.S. diploma	54	27.1
	B.A./B.S.	68	34.2
	Masters/Law	62	31.2
	Dr.	15	7.5

Table 2  
*Demographic Information of Superintendents*

	Demographic Variable	Frequency	Percentage
Gender	Male	170	85.4
	Female	29	14.6
Service	≤ 2 years experience	70	35.2
	3-5 years experience	61	30.7
	6-9 years experience	35	17.5
	≥ 10 years experience	33	16.6

Table 3  
*Demographic Information of School Districts*

	Demographic Variable	Frequency	Percentage
Type	Urban	29	14.6
	Suburban	96	48.2
	Rural	74	37.2
Size	< 2000	74	37.2
	2000-4000	55	27.6
	> 4000	70	35.2

for the interaction between the real and ideal dimensions of leader behaviors based upon the demographic variables of gender, educational level, and years of experience. A 2 (gender) x 4 (years of experience) x 4 (educational level) factorial design was used in examining the independent variables related to board members. A 2 (gender) x 4 (years of experience) factorial design was used in examining the independent variables related to superintendents. The use of MANOVA as the statistical procedure was determined as a result of examining the relationships of groups with two dependent variables (real and ideal scores on the subscales). The results of the study indicated that there were significant differences of superintendents' leader behavior in the perceptions of board members. These differences were observed specifically in the following three areas: there was a main effect of board members' experience on Real Initiation of Structure scores; there was an interaction effect of board members' gender, experience, and educational level on the Ideal Consideration scores; and there was a main effect of superintendents' experience on the Real Consideration scores.

The results indicate that the strongest conclusions from this study are that board members with a high level of experience may perceive their superintendents more positively than do those board members with a low level of board experience on the construct of Initiation of Structure. Univariate analysis of variance (ANOVA) was also conducted on each dependent variable significantly affected by independent variable(s) as a follow-up test to MANOVA. The significance level for ANOVA was set at .05 since when two dependent variables are analyzed, the overall significance level is to be divided by the number of dependent variables being tested (Mertler & Vannatta, 2005). ANOVA results indicate that the real score significantly differs for years of experience ( $F(3, 169) = 3.545, p < .05$ ). Results also indicate that board members perceive their superintendent more positively on the construct of Consideration when a high level of superintendent experience exists. MANOVA results indicate that the main effect of experience (Pillai's Trace = .059,  $F(6, 382) = 1.920, p < .10$ ) had a significant effect on the dependent variables of ideal and real scores of consideration.

Univariate ANOVA and Bonferroni post hoc tests were conducted as follow-up analyses. The effect size was calculated to determine the magnitude of the difference between the groups (Cohen, 1988; Salkind, 2004). ANOVA results indicate that Real Consideration scores significantly differ for superintendent's experience ( $F(3, 191) = 2.737, p < .05$ ), while the Ideal Consideration scores reveal no significant difference ( $F(3, 191) = 1.503, p = .215$ ). Bonferroni post hoc tests for the Real Consideration score indicate that individuals with 3 to 5 years of experience differ significantly from those with 6 to 9 years of experience ( $d = .5956$ ), and those with 6 to 9 years of experience differ significantly from those with 10 or more years ( $d = .6927$ ).

Additionally, the data suggests that female board members hold higher expectations on the construct of Consideration and perceive superintendents' actual behavior lower on Consideration when compared with male board members. The test used for investigating the hypothesis was a two-way multivariate analysis of variance (MANOVA) to determine the effect of gender, length of service, and educational level on the two dependent variables of participants' consideration scores on the LBDQ-Form XII both real and ideal.

Results indicated that the interactive effect of gender, experience, and educational level (independent variables) had significant differences on the dependent variables (Pillai's Trace = .122,  $F(14, 338) = 1.569, p < .10$ ). Individual F tests were performed as a follow-up to MANOVA (Stevens, 1996) indicating significant interaction of board members' gender, experience, and educational level on the Ideal Consideration scores ( $F = 2.571, p < .10$ ) and a significant effect of board members' experience on the Real Consideration score ( $F = 2.497, p < .10$ ). Two-way profiles were analyzed as a follow-

---

up to the MANOVA (Stevens, 1996) to determine differences for the interaction effect of board members' gender, experience, and educational level on the dependent variable of ideal mean consideration scores.

## Discussion

It is evident that the experience levels of both board members and superintendents have the potential to impact perceptions of board members in regards to superintendents' leader behaviors, both ideal and actual (measures as "real"). The data suggests that the more experience board members obtain, the more likely it is that board members will perceive their superintendent's actual leadership behaviors as positive. The finding holds for superintendent experience as well—that is, the more experienced superintendent (as measured in years in the position) is rated more positively than a less experienced superintendent. Evidence also suggests that gender plays a role in board members' expectations and perceptions of superintendents in the area of Consideration: female board members hold higher expectations, rating actual behavior lower than their male counterparts.

The current study extends the knowledge of superintendents' leader behavior as perceived by boards of education. It is evident by the results of this study that board members who have a high level of experience perceive the actual production (Real Initiation of Structure) of their superintendents as significantly higher than those board members with little experience. Additionally, superintendents with 10 years or more of experience are generally perceived more positively than those with less experience regarding their actual concern for people (Real Consideration). Board members who have 10 years or more of experience rate their superintendents higher in the area of Real Consideration.

These results indicate that as experience on the part of boards and/or superintendents grow, board of education members will view the superintendents' actual behavior more positively. The results also suggest that differences in gender account for variability in board perceptions. Female board members who responded to the survey generally hold higher expectations of superintendents in the area of consideration of people (Ideal Consideration) than do males, while the female board members tend to rate superintendents' Actual Consideration lower than do male board members. The results concerning gender provide the insight that female board members may hold higher expectations of superintendents' people skills, while generally rating the actual behavior as lower compared to male board members.

## Implications

Several implications and practical applications are evident as a result of the findings of the study. The role that increased experience plays, both for superintendents and

boards of education, is evident, as well as some gender disparities. The following section discusses the theoretical and practical applications this research engenders.

### *Experience of Board Members*

There is a stark contrast between the training required of superintendents as compared to that of board members. While superintendents generally either come to the position with several other administrative jobs or experiences in their past, any person can be elected to a board of education, and in most of the United States there is absolutely no training required once elected. Additionally, there remains a paucity of available training opportunities if a board of education member so desires such training. An individual may fill a seat on the board with little or no knowledge of the school district's mission, organizational programs, district financial condition, state funding laws, time commitment, governance responsibilities, and administrative and board roles and functions. This lack of knowledge often can be detrimental and may require a great deal of time to acquire the pertinent information necessary to make informed decisions. Moreover, the development of this knowledge base may take months or even years given that this kind of learning often occurs through a variety of real-life situations and a great deal of communication with more experienced board members and the superintendent.

Furthermore, depending on the experience of the board, the superintendent often is placed in the potentially awkward position of training the very board members who will in turn evaluate that same superintendent. As noted in the literature, a school superintendent is placed in a unique employment condition (Cambron-McCabe, Cunningham, Harvey, & Koff, 2005), being hired by, reporting to, and evaluated by the collective group known as the school board. Superintendents often paradoxically spend much of their time in discussion with inexperienced board members assisting them in understanding the roles and functions of boards. This "training" of inexperienced board members usually requires prolonged and conscientious attention, and the superintendent is usually held responsible for this development (Cambron-McCabe, Cunningham, Harvey, & Koff, 2005; Houston & Eadie, 2005).

Related to issues of the experience of board members is a tension that inexperienced board members often do not grasp—that of supporting and governing. In the role of support for the district, board members attempt to ensure the success of the school district, by placing tax issues on the ballot, acting as ambassadors of the organization, and often bringing some level of specific expertise and authority to the district. Delegating to the superintendent responsibility for day-to-day administration of the district, and then acting on the superintendent's recommendations are acts of support. The board also can be supportive by developing a clear job description and setting unmistakable performance expectations of the superintendent.

---

The governance function of the board includes protecting the public interest through selecting the superintendent and treasurer, assessing the performance of these two individuals, setting policy that ensures quality education, and evaluating the district's work. Additionally, maintaining fiscal responsibility, monitoring progress of the strategic plan, and providing oversight for the district's goals are acts of governance. Inexperienced board members often mistake governance for close supervision and end up meddling in minor administrative affairs. Due to their lack of familiarity with the field of education, such meddling can become burdensome for district office personnel and potentially damaging for long-term working relationships. Superintendents and board members would be wise to remember that lack of experience may result in less positive perceptions of the superintendent, and may result in a strained rapport, connections, and associations between the board of education and superintendent. Information from this study can be helpful to both superintendents and board members as they reflect on their relationships in terms of actual leader behavior. Additionally, results lead to the conclusion that there is a tremendous need for board development programs, as well as joint training for both superintendents and boards regarding the roles and functions of each.

### *Gender of Board Members*

The role that gender plays in the expected and actual behavior of superintendents is both interesting, and potentially important, especially when combined with the fact that there remains a scarcity of female superintendents (15% in the current study), while there is more balance with board members (42% are females in the study). These percentages are congruent with those across the state in which the study was completed, with 18% of state superintendents being female, while 33% of state board members are female. This difference in expectations and actual behavior was noticed in the area of Consideration in the current study and may be important for superintendents to clarify expectations in this area to avoid potential conflict with their boards. The possibility exists that there are some inherent problems of understanding the expectations in the area of consideration of people. Perhaps male superintendents tend to be more focused on task and production issues, whereas many board members are focused on people skills of their superintendent. Further, female board members may communicate their expectations differently than do male board members (Borisoff & Merrill, 1985) to primarily male superintendents. In any case, it is a prudent step for a superintendent to seek clarity of expectations from the board of education in this area, especially given the fact that a superintendent's tenure demands that expectations of the board are satisfied (Cambron-McCabe, Cunningham, Harvey, & Koff, 2005).

### *Communication*

Results of the study indicate several points of consideration in the area of communication (note: communication skills and behaviors are measured within both the consideration and initiation of structure sub-scales). Communication must be timely, consistent, and attentive to the needs and expectations of both the board members and the superintendent (Rickabaugh & Kremer, 1997). Communication is critical to an effective, positive perception. The results of this study indicate that more time may be needed in the area of communication with female board members or those who are relatively inexperienced as board members. Awareness of the results of this study may influence the manner in which one communicates with experienced versus non-experienced board members, males versus females.

When communicating with boards of education, experience and observation suggest that practical behaviors on the part of superintendents may enhance the relationship with members of the board. While it may be tempting to emulate and copy another's style, the nature of frequent and often detailed communication with board members should dictate the fact that it is vital to be yourself or one runs the risk of being inconsistent and being perceived as insincere. A second important behavior is to be credible and honest in communication. Superintendents can get themselves into trouble with their boards when attempting to answer every question and appear knowledgeable in every conceivable manner. It is more important to be able to back any statements or answers with factual information, and this often includes going to others for information prior to answering questions. In this way, leaders are more likely to be perceived as effective.

### *Conclusions*

An abundance of prior research has focused and built upon studies of leader behaviors that include a task or production orientation, and one of interaction with people or consideration (Vechio, 2006). These two constructs have been formally defined as Initiation of Structure and Consideration and are measured on the LBDQ as separate subscales and were utilized within this study. In practice, application of these theoretical constructs is evidenced when superintendents and boards work together, addressing the concerns and issues their district faces. During these interactions, public school board members form perceptions of their superintendents, at least in part, based upon superintendents' exhibited behaviors, and these perceptions are critical to the board-superintendent relationship. Boards typically desire a superintendent who is able to "produce results" for the district (e.g.: high state report card scores, increased graduation rates, fiscally responsible management, and other task-related behaviors). Additionally, boards yearn for a superintendent who pays attention to people as individuals, forms positive relationships with parents and community members, and



---

generally works cooperatively with others. As a result of investigating these expectations and perceived behaviors that board members hold, both board members and superintendents have an opportunity for increased understanding and practice of their working relationship.

## References

- Bass, B. M. (1990). *Bass and Stogdill's handbook of leadership* (3<sup>rd</sup> ed.). New York: Free Press.
- Bass, B. M. (1999). Two decades of research and development in transformational leadership. *European Journal of Work and Organizational Psychology*, 8(1), 9-32.
- Borisoff, D., & Merrill, L. (1985). *The power to communicate: Gender differences as barriers*. Prospect Heights, IL: Waveland Press.
- Bryman, A. I. (1986). *Leadership and organizations*. London: Routledge.
- Cambron-McCabe, N., Cunningham, L. L., Harvey, J., & Koff, R. H. (2005). *The superintendent's fieldbook: A guide for leaders of learning*. Thousand Oaks, CA: Corwin Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Conger, J. A. (2004). Developing leadership capacity: What's inside the black box? *Academy of Management Executive*, 18(3), 136-139.
- Dillon, R. R., & Halliwell, J. W. (1991). Superintendents' and school board presidents' perceptions of the purposes, strengths, and weaknesses of formal superintendent evaluations. *Journal of School Leadership*, 1, 328-337.
- Fleishman, E. A. (1973). Twenty years of consideration and structure. In E. A. Fleishman & J. G. Hunt (Eds.), *Current development in the study of leadership* (pp. 1-40). Carbondale, IL: Southern Illinois University Press.
- Fleishman, E. A. (1995). Consideration and structure: Another look at their role in leadership research. In F. Dansereau & F. J. Yammarino (Eds.), *Leadership: The multiple-level approaches* (pp. 51-60). Stamford, CT: JAI Press.
- Goodman, R. H., & Zimmerman, W. G. (2000). *Thinking differently: Recommendations for 21<sup>st</sup> century school board/superintendent leadership, governance, and teamwork for high student achievement*. Published paper by Educational Research Service.
- Gronn, P. (2003). *The new work of educational leaders: Changing leadership practice in an era of school reform*. London: Sage/Paul Chapman.
- Hargreaves, A., & Fink, D. (2006). *Sustainable leadership*. San Francisco: Jossey-Bass.
- Harrington-Lueker, D. (2002). Superintendent rookies. *The School Administrator Web Edition*. Retrieved February 5, 2005, from www.aasa.org
- Houston, P., & Eadie, D. (2005). *The board-savvy superintendent*. Lanham, MD: Rowman & Littlefield Education.
- Hoyle, J. R., & Skrla, L. (1999). The politics of superintendent evaluation. *Journal of Personnel Evaluation in Education*, 13, 405-419.
- Judge, T. A., Piccolo, R. F., & Ilies, R. (2004). The forgotten ones: The validity of consideration and initiating structure in leadership research. *Journal of Applied Psychology*, 89(1), 36-51.
- Kotter, J. (1995). *Leading change*. Boston: Harvard Business School Press.
- Kowalski, T. J. (1999). *The school superintendent*. Upper Saddle, NJ: Prentice-Hall.
- Louis, K. S. (1994). Beyond managed change. *School Effectiveness and School Improvement*, 5(1).
- Mertler, C. A., & Vannatta, R. A. (2005). *Advanced and multivariate statistical methods*. Glendale, CA: Pyrczak.
- Ogawa, R. T., & Bossert, S. T. (1995). Leadership as an organizational quality. *Educational Administration Quarterly*, 31, 224-243.
- Organ, D. W. (Ed.). (1996, May-June). *Leadership: The great man theory revisited*.
- Peterson, G. J., & Short, P. M. (2001). The school board president's perception of the district superintendent: Applying the lenses of social influence and social style. *Education Administration Quarterly*, 37, 533-570.
- Rickabaugh, J. R., & Kremer, M. L. (1997). Six habits to make you a hit with your school board. *The School Administrator Web Edition*. Retrieved July 9, 2006, from www.aasa.org
- Rowan, B. (2002). Rationality and reality in organizational management: Using the coupling metaphor to understand educational (and other) organizations. *Journal of Educational Administration*, 40(6), 604-611.
- Salkind, N. J. (2004). *Statistics for people who (think they) hate statistics*. Thousand Oaks, CA: Sage.
- Spillane, J. P. (2006). *Distributed leadership*. San Francisco: Jossey-Bass.
- Stogdill, R. (1963). *Manual for the leader behavior description questionnaire, Form XII*. Columbus, OH: The Ohio State University, Bureau of Business Research.
- Tallerico, M. (1989). The dynamics of superintendent-board relationships: A continuing challenge. *Urban Education*, 24, 215-232.
- Vechio, R. P. (2006). *Organizational behavior: Core concepts*. Mason, OH: Southwestern.
- Weick, K. E. (2001). *Making sense of the organization*. MA: Blackwell Business Press.
- Yukl, G. A. (2002). *Leadership in organizations*. Englewood Cliffs, NJ: Prentice Hall.

---

# *An Exploratory Survey of Participants in Urban and Suburban Teacher Academy Programs*

Leah Wasburn-Moses  
Molly Kelly-Elliott  
Miami University

## **Abstract**

*In an era plagued by teacher shortages, recruitment and training strategies are crucial. One promising early recruitment strategy is the use of teacher academies, pre-collegiate programs for aspiring teachers. Yet precious little is known about these programs or their students, many of whom appear to mirror the much sought-after diversity hoped for in practicing teachers (Darling-Hammond, Berry, Hasselkorn, & Fideler, 1999). This paper reports results of a survey of 133 students enrolled in 11 teacher academies surrounding a large Midwestern city. Data on students' demographics, future goals, and perceived program needs are reported. Results indicate strengths in program experiences and cohesiveness, and the need for greater emphasis on college readiness and program organization. Further research on these programs is imperative.*

## **Introduction**

Education professionals and critics alike agree that there is a critical shortage of high quality teachers in this country (Ingersoll, 2004; Levine, 2006). In particular, there is a need for increasing the number of teachers from diverse backgrounds in order to reflect the growing diversity of our nation's population (Darling-Hammond & Bransford, 2005). Although such shortages are often attributed to the "graying" of the teacher workforce, many forces are at play, including rapid turnover, teacher misassignment, low entry standards, and the low status of the teaching profession in society (Ingersoll, 2004). Many strategies have been employed in order to increase teacher recruitment. Some of these strategies include alternative certification programs, community college programs, financial incentives, and precollegiate clubs and academic programs (Wilson, Bell, Galosy, & Shouse, 2004). Darling-Hammond, Berry, Hasselkorn, & Fideler (1999) outline five "leading-edge recruitment efforts," which include precollegiate initiatives, traditional university-based programs, community college pathways, paraprofessional pathways, and post-baccalaureate alternative certification programs.

Perhaps the least researched of these recruitment efforts is precollegiate teacher recruitment. Although this strategy encompasses clubs, mentoring programs, and stand-alone internships, full-fledged precollegiate teacher recruitment programs are referred to as "teacher cadet programs," "teaching career academies," and "teacher academies" (Wilson et al., 2004). Participants in these programs "are more likely to demonstrate persistence into (the) teaching profession than less intensive approaches" (Berrigan & Schwartz, 2000, p. 7). The official purpose of teacher academies is to "nurture and 'grow' prospective teachers committed to serving their schools and communities" (Berrigan & Schwartz, 2000, p. 8). Such programs have been described as "teaching-focused,

comprehensive academic programs within larger schools" (Berrigan & Schwartz, 2000, p. 8). In 1995, a national survey uncovered the existence of 253 teacher academies involving over 50,000 potential teachers. Further, a full 64% of the participants were young people of color (Darling-Hammond et al., 1999). This finding that precollegiate teacher recruitment programs attract disproportionate numbers of minority students has been replicated in several other studies (Berrigan & Schwartz, 2000). Although most of the better known precollegiate teaching programs have been created as a result of state-level initiatives, many others are locally conceived and implemented (Clewell, Darke, Davis-Googe, Forcier, & Manes, 2000).

Although teacher academy programs are quite diverse, overall they tend to feature "teaching, tutoring, and mentoring experiences in a variety of settings" (Darling-Hammond et al., 1999, p. 201). They are typically contained in the last one to two years of high school. Berrigan and Schwartz (2000) explain that most teacher academy programs consist of three major components: "electives related to teaching, learning, and children; pre-college internships at local elementary, middle, and high schools; and partnerships with colleges/universities that provide a 'pathway' or corridor into college and teacher education" (p. 3). Typical course content is described as "learning theory, classroom management, multiculturalism, child development, and assessment" (p. 3).

Advantages to the student include offering positive images of the teaching profession; a "rewarding sense of responsibility and a powerful connection to the children they work with" (p. 2); specialized courses that present material similar to that covered in college coursework, but in a creative and appealing manner; and an opportunity to clarify their career goals (Berrigan & Schwartz, 2000). Other reported benefits include a better understanding of the teaching process and its complexities, as well as better preparation for college in general (Darling-Hammond et al., 1999).

---

Due to the growing popularity of these programs, in 2000 Berrigan and Schwartz produced a guidebook for the development of teacher academy programs, entitled "Urban Teacher Academy Project Toolkit: A Guide to High School Teaching Career Academies." The publication was sponsored by the United States Department of Education's Office of Vocational and Adult Education. This booklet presents guidelines for program planning and implementation, including how to locate resources, how to create recruitment and retention policies, and how to design courses and internships.

Although teacher academy programs have precipitated some interest in the professional literature, unfortunately, few quality evaluation studies have resulted. The South Carolina Teacher Cadet Program is the only teacher academy program with sufficient scope and longevity to produce data that is even somewhat reliable (Darling-Hammond et al., 1999; Southeast Center for Teaching Quality, 2003). This program began during the 1985-1986 school year, and currently serves over 2200 students each year in approximately 150 high schools throughout the state. The cadet program is supported by 22 of the 30 colleges and universities with teacher education programs, in that they grant college credit for program completion. Outcome data indicates that 35% of program participants report an intent to pursue teaching careers.

Clearly, the question of whether and how these programs contribute to the amelioration of teacher shortages and to the improvement of teacher quality is of utmost importance, and goes beyond intent of graduates to teach. These programs originally grew out of research indicating that individuals make career choices much younger than previously thought, and that those entering the teaching profession are particularly influenced by their own experiences as K-12 students (Darling-Hammond et al., 1999). Although research has not shown a clear connection between the use of these programs and the lessening of teacher shortages, we can turn to the large-scale evaluation studies conducted on the South Carolina Teacher Cadet Program. Of the former program participants in 1987-1988 who could be traced, approximately two-thirds were teaching in the state in 1993. About 30% of these were teaching in rural schools, and 29% in what was considered a "critical shortage area" (Darling-Hammond et al., 1999, p. 203). This information appears to be the only existing data regarding the effect of such programs on teacher shortages. However, due to flaws in data collection and large attrition rates, few other claims about impact on teacher shortages, and no claims regarding teacher quality can be made (Darling-Hammond et al., 1999).

Regardless of the gaps in the existing data, Wilson et al. (2004) conclude their review of strategies to increase teacher supply with the statement that among the "most promising minority recruitment programs" include locally-developed programs designed to encourage students of the community to become teachers and to return to their home communities to teach. Yet a thorough review of the literature uncovered almost no data-based studies on precollegiate teacher recruitment in the past decade.

The current study is an investigation of the background and perceptions of some of the participants enrolled in teacher academy programs in 11 high schools surrounding a large Midwestern city. Due to the need for knowledge about the nature of these teacher recruitment efforts, three research questions were addressed:

1. How do student demographics compare to the demographics of practicing teachers?
2. What are students' perceptions of the strengths and weaknesses of their teacher academy programs?
3. What are students' perceptions of their current and future needs?

## Method

A survey instrument was administered to 133 students in 11 "teacher academy programs" at three urban, seven suburban, and one rural high school surrounding a large Midwestern city. The students were participants in one of two one-day university-based programs. Quantitative analysis, in the primary form of descriptive statistics, was conducted on participant and program demographics. Qualitative analysis was conducted on participants' responses to open-ended survey questions.

### *Instrumentation*

Due to the lack of data on these particular programs and their students (Darling-Hammond et al., 1999), a new survey instrument was designed. It was not based on previous research. The instrument addressed student demographics, perceptions of programs, and perceptions of current needs. Demographics collected included race and sex, as well as the type of teaching career participants planned to pursue (i.e. grade and/or subject). Participants also were asked to report strengths and weaknesses of their teacher academy programs. Finally, the survey instrument included three questions addressing current needs: (1) What skills do you think you need to improve before you start college?, (2) What else do you think you need to know before becoming a teacher?, and (3) How will you choose what college to attend?

### *Data Collection*

Data were collected at two points in time, one in 2007 ( $n = 48$ ) and one in 2008 ( $n = 85$ ). Participants were involved in one of two one-day symposia for teacher academy students whose schools fell within a one-hour driving radius of the hosting institution. The workshop was housed at a public university and designed as a recruitment tool to entice future teacher candidates to attend the university after their graduation from high school. The symposium consisted of various presentations and activities designed by teacher academy and university faculty.

Participants were teacher academy students who had elected to participate in the symposium. All students were

requested verbally to participate in the study by filling out the one-page survey at some point during the symposium. The survey instrument was included in students' packet of materials from the host institution. Surveys were collected by hand at the conclusion of the symposium.

### Participants

Participants were 133 high school students, 44 male (33%) and 89 female (67%), from 11 teacher academy programs at high schools surrounding a large Midwestern city. Although the first teacher academy program in the geographic area began in 1991, the majority of the programs included in the current study were less than six years old. Although the majority are two-year programs, some are one year and one is a four-year program. All are two class periods per day. Students in each of the 11 programs are required to take a certain number of teaching-related courses and participate in a set of field experiences, much like student teaching. These experiences culminate in creation of a final portfolio project, many of which are aligned to Praxis standards. Portfolios are scored by a local teacher and by a university professor. Many of the teacher academies in this area have articulation agreements with universities in the region, whereby graduates can acquire university credit for successful completion of the teacher academy program (R. Terrell, personal communication, April 18, 2008).

The data on student gender is relatively comparable to state-level data on the gender of practicing teachers, showing that 29% are male, and 71% female. The majority (64%) of the participants described themselves as "White," which can be compared to data indicating that 93% of the state's teachers consider themselves "Non Minority" (Ohio Department of Education, 2007b). About one-third of the participating students identified themselves as "Black" (33%), one student self-identified as "Hispanic," and three students as "Other" (3% combined). State-level data show that approximately 7% of the state's teachers are Black, and a negligible number self-identify in other racial categories (Ohio Department of Education, 2007b). Future teaching preferences revealed that 38% of the participants planned to pursue elementary teaching careers, 50% planned for secondary careers, 8% indicated a desire to pursue a specialist area (speech/language pathologist or intervention specialist) and 4% selected early childhood education (preschool or kindergarten).

Of the 11 Teacher Academy programs, three were in urban school districts, seven programs were in suburban school districts, and one was in a rural district. Demographic data of participating schools are reported in Table 1. Most of the schools had a Caucasian majority, although percentages ranged from 4% to 93%. Rates of economic disadvantage (as indicated by percent of students who qualified for free or reduced lunch) ranged from 4% to 64%. The ESL population ranged from none to four percent, and the schools' graduation rates ranged from 77% to 99% (Ohio Department of Education, 2007a).

Table 1  
*School Demographics*

High School	Location	Percent Caucasian	Percent FRL*	Percent ESL	Graduation Rate
1	suburban	81%	16%	1%	89%
2	suburban	82%	7%	2%	99%
3	urban	82%	35%	4%	77%
4	urban	4%	64%	2%	91%
5	suburban	84%	7%	3%	95%
6	suburban	24%	52%	0%	89%
7	urban	74%	46%	3%	85%
8	suburban	37%	35%	3%	96%
9	rural	94%	7%	0%	91%
10	suburban	82%	4%	1%	96%
11	suburban	18%	28%	2%	87%

\* Free & Reduced Lunch

### Data Analysis

Quantitative data in the form of demographics were analyzed through simple descriptive statistics. Qualitative data were analyzed through inductive analysis (Janesick, 2000). Selective coding was used to create categories encompassing each response. Codes were compared to subsequent responses and revised in order to continue to encompass each response. Later, broader categories were developed from the codes in order to develop a reduced list that encompassed the majority of responses. "Other" categories were created to ensure that all responses were included (Charmaz, 2000). As data were added, the categories were revised as needed to continue to encompass participant responses.

## Results

### Program Strengths and Weaknesses

First, participants were asked "What do you think are the strengths and weaknesses of your teacher academy program?" As shown in Table 2, four categories emerged from responses to this question. The themes included positive experiences in the program, affective characteristics, an emphasis on Praxis domains, and a focus on writing/organizing. Nearly half (45.1%) of the respondents indicated that program experiences were a strength of the program, referring to hands-on internships, practica, and tutoring experiences. They used words such as "the amazing experiences we have," "multiple outside experiences," and "real world experience in internships and field experiences" to describe program strengths. Nearly half (42.1%) also rated affective characteristics as a strength. These responses appeared to emphasize the emotional closeness of the students and teachers in the teacher academy. For example, "we're basically like a family," "our teacher is amazing," "all of us are very close and supportive of each other," and "we are all very close and help each other to achieve our goals." Minor categories included knowledge of Praxis domains (6.8%), citing "domain review," "Praxis exposure," "learning the domains," and "well educated on the domains" as strengths of the program. Even fewer (3.0%)

respondents described writing and organizing as strengths, with comments such as “organizing and understanding,” and “we get ahead of the game (in writing).” Finally, just under ten percent (9.8%) of respondents listed strengths categorized as “other.” These included generic comments that did not point to any specific strengths, such as “we get a lot of knowledge before college,” “prepares you for the future,” “shows us a wide range of options in education,” and “make sure I know I want to teach.”

Table 2  
*Perceived Program Strengths and Weaknesses*

Strengths*		Weaknesses	
Experiences	45.1%	No Weaknesses	23.8%
Affective Characteristics	42.1%	Behavior Issues	22.2%
Praxis Domains	6.8%	Insufficient Skills	16.7%
Writing / Organizing	3.0%	Insufficient Time for Program	13.5%
Other	9.8%	Place of Program in School	6.3%
		Other	17.5%

\*Participants provided more than one response

Weaknesses presented by participants were organized into six categories: no weaknesses, behavioral issues in the classroom, insufficient skills, insufficient time for the program, the place of the program in the school, and other. The largest response category was “none,” with 23.6% of respondents indicating no weaknesses in the program. This question was either left blank when all other questions were filled out, or answered with “none” or “n/a.” Over one-fifth (22.2%) indicated that behavioral issues in the classroom were a concern, with comments such as “do not listen to each other,” “situation resolvment,” “due to being close we fight a lot,” “kids not getting along,” “keeping authority,” and “not staying on task.” Fewer of the respondents (16.7%) indicated a need for more skills as a weakness of the program. These skills included general academic/college preparation skills as well as specific skills for teaching children. Comments included “not enough about Praxis,” “getting in depth on subjects/lessons,” and “focus on child behavior.” Insufficient time for the teacher academy program was a theme in 16.7% of responses. For example, participants noted “lack of time,” “classroom time—little,” and “not enough time in the day!” Although a minor theme, 6.3% of respondents described place of program in school as a weakness, in that programs were new, under enrolled, and/or experience a lack of recognition in the school. Representative statements included “our numbers are too low,” “we need more people to be interested and join,” “not recognized in school; needs more interest,” and “understaffed.” Finally, 17.5% of the respondents listed weaknesses categorized as “other.” These diverse comments addressed their teacher’s personal characteristics or specific classroom expectations, as well as thought fragments (e.g. we spend time on).

### *Skills Needed*

Second, participants were asked, “What skills do you think you need to improve before you start college?” Four categories encompassed these responses, which are shown in Table 3. The majority (63.3%) of respondents indicated a need for improved organization/study skills, in that they felt they needed to learn more about managing time, both inside and outside class. Representative comments included “I need to improve my time management skills,” “to not wait until the last minute and jump ahead,” “my organization and procrastination skills—work on meeting deadline,” “scheduling, study/research skills, time management,” and “studying skills and completing homework.” Fewer (13.3%) respondents indicated a need for improved academic writing abilities with statements such as “writing and reading skills,” “writing more developed papers,” and “writing skills and expression.” Just over ten percent (10.2%) of respondents mentioned a need for improving their own communication and/or social skills. Representative statements included “I need to control my temper more,” “attitude readjustment,” “getting more comfortable meeting new people,” and “public speaking skills.” Finally, 13.3% of the respondents noted needed skills categorized as “other.” These skills primarily encompassed academic skills in the content areas (e.g. “building foreign language skills” and “I think I would need to improve my math skills,” and specific skills related to teaching (e.g. “memorizing my domains more” and “improve on my lesson planning”).

Table 3  
*Skills to Improve*

Organization / Study Skills	63.3%
Writing	13.3%
Communication / Social Skills	10.2%
Other	13.3%

### *Additional Information Needed*

Third, participants were asked to respond to the question, “What else do you think you need to know before becoming a teacher?” Many of the responses to this question were too diverse to be organized into categories, as shown in Table 4. For example, 30.7% indicated multiple needs in their responses to this question. These responses included comments such as “a lot,” “everything—you can never stop learning about being a good educator,” “many different aspects of the classroom,” and “pretty much everything.” As such, they were difficult to categorize. About one-fifth (20.2) indicated a need for more knowledge about teaching methods and planning. These responses encompassed curriculum and instruction, and the knowledge behind what to use when. Representative statements included “content standards,” “necessary curriculum,” “how to get every one of your students to understand what you teach and be helpful in more different ways,” and “connecting lesson plans and activities to the info being taught.” Just over 10% (10.5%) indicated

a need for more knowledge about child development and students in general, with comments such as “I would like to learn more about children’s development,” “more about child development,” and “to learn more about children.” Fewer (15.8%) of the respondents indicated a need for either general or specific knowledge in classroom management, with comments such as “just how to handle my own classroom,” “more discipline techniques,” and “I need to know how to draw the line between friend and authority.”

Table 4  
*Additional Skills Needed*

Multiple Needs	30.7%
Teaching Methods/Planning	20.2%
Classroom Management	15.8%
Students/Child Development	10.5%
Other	22.8%

However, over one fifth (22.8%) of responses to this question were too diverse to be categorized. These statements included generic comments about the job market in teaching (e.g. “I want to know the best places in the country to teach” and “how bountiful is the job market”), personal questions about best fit of college major (e.g. “whether I can minor in gifted education” and “which subject I would best teach”), and generic comments not related to any one specific skill, such as “what we need to succeed” and “just the basics”).

Finally, participants were asked, “What additional information do you need to help you choose the college you will attend?” Responses were organized into four categories, as shown in Table 5. Almost half (48.4%) of the respondents indicated that campus offerings and programs were important to their college decision. These offerings and programs included majors, courses, and extracurriculars, as well as general issues of quality of teacher education programs and of the institution overall. Comments included “programs offered,” “more class info,” “How good their music education program is,” “if I can double major in History Education and Latin Education,” “I would like to know if the college meets my needs academically,” “I would want to know if the college had a good teaching program,” and “the majors and some information of that college.”

Table 5  
*Additional Information for College*

Program/Campus Offerings	48.4%
Already Decided	24.6%
Monetary Issues	18.0%
No Information Needed	9.0%

About one-fourth (24.6%) of the respondents indicated that they had already made their college decision. These remarks included “I have already decided,” “none, already applied and accepted,” and “I will be attending X University in the fall.” Just under one-fifth (18.0%) of the respondents indicated that monetary issues were important to their de-

cision. These issues included availability of scholarships, financial aid, and on- and off-campus jobs. Representative comments included “The financial aid and programs offered,” “what I need to do to be financially stable,” and “education costs and availability of scholarships.” Finally, just under 10 percent (9%) replied that no other information was needed, by responding “none,” “nothing,” or “N/A.”

## Discussion and Implications

Several themes emerged from the results of this study that point to general strengths and needs of the teacher academy programs included in this study. First, student demographics do point to positive implications for teacher academy programs as a potential recruitment tool for teachers from more diverse populations. Second, strengths appear to be in the areas of practical experiences offered to students and in affective areas such as cohesiveness of program (e.g. feelings of belonging, enthusiasm for teaching). These strengths can be seen as crucial assets to the field of teacher preparation. Third, various program issues point to the possible need for goal setting. Needs appear to be in two main areas: an increased focus on college readiness skills, and increased attention to teaching-related knowledge. Future research should involve content analysis of teacher academy programs and their partners, and in-depth, longitudinal studies of teacher academy participants.

First, although a very small sample from a limited part of one state, the demographics of student participants do look promising. Students show more racial diversity than do practicing teachers in the state, and they are at very least being exposed to the teaching profession, as well as experiencing a college preparatory program. These findings support those from previous studies, even those with large samples (Berrigan & Schwartz, 2000; Darling-Hammond et al., 1999).

Second, students mentioned repeatedly the importance that practical experience was playing in their program, referring to hands-on, practical internships, field experience, tutoring, etc. This finding also supports previous research indicating the major strengths of teacher academies are in the areas of offering practical experiences that show teaching in a positive light (Berrigan & Schwartz, 2000). A strength in practical experience is particularly advantageous to future teachers, as critics have long called for increasing number and duration of supervised field experience in all stages of teacher education (Darling-Hammond & Bransford, 2005). Perhaps traditional teacher preparation programs can learn from the model of the teacher academy.

Another strong aspect of the teacher academy programs, according to students’ reports, appears to be in the affective domain of the program. Students’ responses about their programs, in terms of their uplifting, active nature, and their descriptions of personal relationships with teachers and fellow classmates appeared quite positive. Again, teacher academy programs have been lauded in the past for their ability to raise teaching as a profession and to privilege the positive

---

teacher-student relationship (Berrigan & Schwartz, 2000). The teacher academy programs described by participants appear to be small communities with tight connections. Again, researchers and school leaders have attempted to produce such close communities of teachers for years, believing in their potential to contribute to teacher growth (Ball & Cohen, 1999). This strength offers potential to nurture learning in a variety of ways. It is also possible that these positive affective experiences may be contributing to many participants' reluctance to indicate any weaknesses with their programs.

Third, the needs that emerged from students' responses are also enlightening. Although many of the participants indicated that their teacher academy programs had no weaknesses, the most common response of those who did report a weakness was behavioral issues in the teacher academy classroom. Combined with comments surrounding insufficient time and the place of the teacher academy program in the school at large, it is possible that these programs do not have a well-defined mission or structure. Further, many are new and may not be well established within the larger school context. Perhaps increasing partnerships with other teacher academy and university programs might assist with goal-setting, mentorship, and establishing structure and consistency, as well as legitimacy to outside schools and programs. Similar needs are also reflected in the professional literature. For example, one reported challenge for teacher academy programs is setting and communicating clear programmatic goals (Berrigan & Schwartz, 2000). Perhaps most important is consideration of teacher academy programs as one part of a holistic preparation for college. Examining teacher academy programs from this perspective shows that themes raised by data collected in this study align well with programs created to encourage and prepare students from disadvantaged populations to attend college (Swail & Perna, 2002).

As for specific content needed, several important issues emerged from students' discussion surrounding what skills they still need before becoming a teacher. These needs include organization and study skills. Students referred to several skills that have been determined to be crucial for college-level work, such as time management, independent study skills, and writing skills. Many students indicated the need for more time in the teacher academy program. It is possible that the balance between presenting skills for teaching and presenting generic study/college preparedness skills is difficult to strike. Again, explicit goal-setting might provide some assistance in this area.

Although practical experiences appear to be a clear strength of these programs, participants' responses indicated more mixed results in the theoretical, or classroom instruction aspect of the teacher academy. For example, when asked what else they needed to know before becoming a teacher, participants' responses varied too widely to categorize. Knowledge for teaching such as classroom management, student characteristics, etc. appears as though it may be overwhelming for the participants in this study, given the number who indicated multiple needs for learning. It is important to note that a teacher academy is precollegiate; it

is not expected to teach everything a teacher needs to know (Berrigan & Schwartz, 2000). This result could simply be attributable to students learning about the sheer complexity of teaching rather than to a deficit in teacher academy programs themselves. Interestingly, a benefit of teacher academies mentioned in the professional literature is giving students a view of teaching as a complex profession (Berrigan & Schwartz, 2000). Therefore, this finding may simply be a reaction to this important aspect of the teacher academy. However, development of priorities or standards for inclusion in coursework of teacher academy programs might not only alleviate some issues regarding limited time, and/or place of the program in the school, but might be one way to narrow content to complement the traditional college programs students hopefully will be entering later.

Future research on this topic is imperative (Darling-Hammond et al., 1999). Experts in teacher recruitment agree that precollegiate recruitment efforts hold much promise for increasing teacher supply and diversity, yet few recent efforts have been made even to locate these programs, much less evaluate either their practical or their theoretical components. First and foremost, program models and characteristics should be documented. Content covered and time spent on various activities and coverage of content is also important knowledge for the field. Perhaps most important, longitudinal research could demonstrate the impact of such programs, in terms of which programs are most successful at producing future teachers, particularly those from diverse backgrounds.

However, it is important to note that even the best of precollegiate programs from a programmatic perspective will be unsuccessful if students are unable to attend or complete college or university programs. The teacher academy programs can be seen as an ideal place to explore and implement what the field of education has learned about the importance of field experiences, mentorship, transition to college, and initiation into the culture of teaching. To date, though, little attention has been paid to these efforts in the research literature. Some important questions for further study include:

1. How do teacher academy programs balance preparation for college with preparation for teaching, and allow high school students to explore teaching as a career option without precluding exploration of other career possibilities?
2. How can the various models of teacher academies be categorized, and how are these models tied to student outcomes (e.g., college entrance, college completion, entrance into teaching)? Do teacher academies have a well-described mission?
3. What support systems (e.g., university partnerships, linkages with other teacher or career academies) can help alleviate some of the issues found in this and previous research?
4. How does the use of teacher academies compare with other forms of teacher recruitment?

---

## Limitations and Conclusion

There are several significant limitations to this study that limit generalizability. The first limitation is self-selection, in that participants involved in this study elected to participate in this specific university-based program; therefore, this sample may not be representative of teacher academy students as a whole. In particular, the sample may have been more positive about their programs and their choice of careers than their classmates who did not choose to attend. Further, the fact that this particular university was attempting to recruit survey participants as students and also giving them a survey might have influenced the way students responded.

Additional limitations include a small sample from a small geographic region. Ideally, participants would be selected at random. Further, individual students' perceptions were not accompanied by direct observation or assessment. There is no triangulation of data reported in this study.

In conclusion, despite the relatively small scope of the current study, teacher academies do appear to warrant further study. These programs do seem to be attracting more diversity and an enthusiastic population of students. Perceived program strengths appear to be in the areas of practical experiences and a sense of program unity. These areas have been found to be crucial to the development of teachers, providing another reason why teacher academies deserve more attention in the professional literature (Darling-Hammond & Bransford, 2005). Perceived needs include more focus on college success skills and knowledge for teaching. However, it is important to note that not all of these skills may be within the scope of the teacher academy.

Goal setting is crucial, particularly in determining what content is most important to present to students at this level. Integration of the teacher academy program within the larger high school academic program should be given further consideration. Further, research is quite weak in this area; some areas of future interest are in identifying and categorizing existing programs and support systems, tying program models to outcomes, and comparing effectiveness of different recruitment models.

## References

- Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In G. Sykes and L. Darling-Hammond (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 3-32). San Francisco: Jossey Bass.
- Berrigan, A., & Schwartz, S. (2000). *Urban teacher academy project toolkit: A guide to high school teaching career academies*. Washington, DC: Office of Vocational and Adult Education.
- Charmaz, K. (2000). Grounded theory: Objectivist and constructivist methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2<sup>nd</sup> ed., pp. 509-336). Thousand Oaks, CA: Sage.
- Clewell, B. C., Darke, K., Davis-Googe, T., Forcier, L., & Manes, S. (2000). *Literature review on teacher recruitment programs*. Washington, DC: U. S. Department of Education.
- Darling-Hammond, L., Berry, B. T., Hasselkorn, D., & Fidler, E. (1999). Teacher recruitment, selection, and induction: Policy influences on the supply and quality of teachers. In G. Sykes (Ed.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 183-233). San Francisco: Jossey-Bass.
- Darling-Hammond, L., & Bransford, J. (2005). *Preparing teachers for a changing world*. San Francisco: Jossey-Bass.
- Ingersoll, G. M. (2004). Four myths about America's teacher quality problem. In D. Miretzky (Ed.), *Developing the teacher workforce: 103<sup>rd</sup> yearbook of the National Society for the Study of Education* (Vol. 1, pp. 1-33). Chicago: University of Chicago.
- Janesick, V. J. (2000). The choreography of qualitative research design: Minuets, improvisations, and crystallization. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2<sup>nd</sup> ed., pp. 379-400). Thousand Oaks, CA: Sage.
- Levine, A. (2006). *Educating school teachers*. Washington, DC: The Education Schools Project.
- Ohio Department of Education. (2007a). Ohio District Report Cards. Retrieved June 6, 2007, from <http://www.ode.state.oh.us>
- Ohio Department of Education. (2007b). Staff Data. Retrieved April 13, 2008, from <http://www.ode.state.oh.us>
- Powell, A. G., Farrar, E., & Cohen, D. K. (1985). *The shopping mall high school: Winners and losers in the educational marketplace*. Boston: Houghton Mifflin.
- Southeast Center for Teaching Quality. (2003). *Recruitment and retention strategies in a regional and national context*. Retrieved March 22, 2008, from <http://www.teachingquality.org>
- Swail, W. S., & Perna, L. W. (2002). Pre-college outreach programs: A national perspective. In W. G. Tierney & L. S. Hagedorn (Eds.), *Increasing access to college: Extending possibilities for all students*. Albany, NY: State University of New York.
- Wilson, S. M., Bell, C., Galosy, J. A., & Shouse, A. W. (2004). "Them that's got shall get": Understanding teacher recruitment, induction, and retention. In D. Miretzky (Ed.), *Developing the teacher workforce: 103<sup>rd</sup> yearbook of the National Society for the Study of Education* (Vol. 1, pp. 145-179). Chicago: The University of Chicago.



---

# *The Effect of Supplemental Educational Services on Student Learning Outcomes*

Jane Beese  
University of Akron

## ***Abstract***

*This study investigated the effects of Supplemental Educational Services on student achievement in reading and math. A meta-analysis was conducted to investigate the relationship between student achievement gains in reading and participation in SES. Six studies were used in the meta-analysis with a total of 395 participants. The combined weighted effect size correlation was  $r = 0.48$ . A meta-analysis was conducted to investigate the student achievement gains in math. Two studies were used in the meta-analysis, with a total of 136 participants. Their combined weighted effect size correlation was  $r = 0.04$ . The magnitude of this effect size correlation is trivial, indicating that participation in math SES produces little effect; however, the effect size correlation associated with reading SES suggests a large effect on learning outcomes in the content area of reading.*

## **Introduction**

In the 20<sup>th</sup> century, educational reform played a central role in the war on poverty. The Elementary and Secondary Education Act of 1965 (ESEA), a Federal law largely defined by Title I, focused on lessening the disparities in academic performance between poor and wealthy schools by promoting equitable academic achievement (Stullich, McCrary, & Roney, 2006). Initially federal funds were used by states to equalize school funding and ensure alike treatment but as time went by with little change and growing concern about international competition, this focus evolved into a standards-based accountability system with the objective of equal outcomes for all students (Bowe, Cronin, Kingsbury, & McCall, 2005).

Two presidential education initiatives have preceded the current efforts of NCLB. The first was Goals 2000: Educate America Act of 1994. Goals 2000 established a framework in which to identify world-class academic standards, to measure student progress, and to provide the support that students may need to meet the standards (Carr, 2001). By 2000, the goals were for all children starting school to be ready to learn, to increase high school graduation rates to 90%, for students in grades 4, 8, and 12, to demonstrate competency in specified subject areas, for all students to be ready to enter the workforce, for students to rank first in the world in math and science, for every school to be free of drugs, and for teachers to have access to professional development. The National Education Goals Report demonstrated modest improvements in several goals, which included more children born with a healthier start in life, more families reading and telling stories to their children, improved math scores for students in fourth and eighth grades, and more degrees earned in math and science. In other areas, the nation regressed; 12<sup>th</sup> grade reading achievement declined, fewer secondary school teachers held degrees in their subject areas, the gap in college completion rates between White and Hispanic students widened, and school violence increased. Other areas showed no change.

In 1994, the US legislature passed a second presidential initiative, the School-to-Work Opportunities Act (STW). This law required states to coordinate school-to-work plans with the educational reforms they were already planning with Goals 2000. Both acts involved a restructuring, rescheduling, and rethinking of educational practices—in other words, a systemic change within education. The failure of STW can be linked to the deficiency in understanding the process of implementing curriculum change or change in traditional subject areas. Suggested improvements were to develop a concentrated effort to influence pre-service teacher education programs within the higher education structure, a definitive body of knowledge, a scope and sequence for the new material, professional development models, an administration involvement plan, and an evaluation plan using subjective normative testing. According to Carr (2001), “Future federal educational efforts should be able to improve from the STW shortfalls and create a more effective design and implementation methodology” (p. 34). It has been suggested that educational reform should be developed from programs that provide information and statistics so that knowledgeable decisions can be made (Kruse, forthcoming).

## *No Child Left Behind*

Historically, the No Child Left Behind Act of 2002 (NCLB) is the largest kindergarten through 12<sup>th</sup>-grade federal education program. NCLB represents a restructuring and redirection of federal efforts to support elementary and secondary education. Prior to NCLB, public school accountability had been a state and local responsibility with the federal government and national organizations playing a supportive role. In the history of education reform, there is no federal law that exceeds the nationalization of education policy such as NCLB (Elmore, 2004). The federal government has become highly involved in the daily operation of public education by instituting a federal law that imposes a single accountability system determined suitable for all schools while setting national parameters on state and local

---

accountability systems (Bowe, Cronin, Kingsbury, & McCall, 2005; Cocoran & Goertz, 2005; Elmore, 2004). As a result of the expanding federal role, NCLB has raised fundamental issues regarding who controls education (Sunderman, Kim, & Orfield, 2005).

Behind the actions taken on behalf of NCLB lie four principles. First, the law provides an accountability system that identifies underperforming schools. A fundamental component of accountability is testing; NCLB put in place a system of testing to validate school effectiveness and holds schools and districts accountable for student learning. Many educators advocate this type of approach; i.e., one that designates a specific and well-planned curriculum and uses tests to determine the extent the curriculum has been taught and learned (Glickman, Gordon, & Ross-Gordon, 2007; Ravitch, 1985; Walberg, 2001).

Second, the law increases the opportunities for students who are in schools that are deficient in the subject areas of reading and math. Two school options given to parents in underperforming schools are school choice and Supplemental Educational Services (SES). Through these options, students are provided with more opportunity to get the help they need.

Third, the law increases the capacity for students to become proficient. NCLB policymakers have set a target rate of 100% proficient by 2014. By improving the overall effectiveness of the public school system, the path is potentially cleared for all students to become proficient.

Fourth, the law attempts to reduce the achievement gap among subgroups of the population. A key feature of NCLB is the goal of narrowing the achievement gap between white and minority students by using a federally-required set of measures. States and school districts must report on the progress of specific subgroups of students (e.g., racial/ethnic, impoverished, disabled, and limited-English proficient).

Under NCLB, every state is required to set standards of academic content and measure each student's yearly progress in the core subjects of reading, math and science. Adequate yearly progress (AYP) is the determinant of school improvement. A school that fails to meet AYP is classified as in "improvement" status. Schools identified for improvement are subject to a series of sanctions. Sanctions are based on markets and privatization theories (Sunderman, Kim, & Orfield, 2005). Schools in the first year of improvement are required to offer to all students the option to transfer to another school not in improvement status. Students in schools in the second year of improvement are eligible for Supplemental Educational Services.

Both public school choice and Supplemental Educational Services are provisions of NCLB intended as corrective action for schools identified as needing improvement. Parents of eligible students are notified by their child's school and may select from a list of qualified providers (Kruse, Liang, & Beese, 2005, 2006; Kruse, Liang, & Widenbaugher, 2004). Students who come from low-income families attending Title

I schools, whether or not they performed at proficiency, are eligible for SES (Corwin & Wilhelm, 2006). In circumstances where only a limited number of students can be provided services due to financial constraints or other limitations, priority is given to the lowest achieving students.

### *Supplemental Educational Services*

Supplemental Educational Services, otherwise known as SES, are educational activities provided outside the normal school day, designed to enhance the educational services that are provided during a regularly scheduled school day (Kim & Sunderman, 2004; Smole, 2006a). In simple terms, Supplemental Educational Services (SES) are after-school tutoring programs. Since 2003, the SES school option has afforded students in failing schools access to tutors where they can get assistance in reading and math. SES providers can vary by type. Providers can be district or school programs or private tutoring businesses. All providers must be approved by each state's Department of Education and aligned with the state's reading and math content standards.

The money to finance SES is provided by redirecting Title I funding. The former goal for Title I was to provide educational opportunities for identified groups of disadvantaged children. This goal has now shifted to individualized access to educational programs. NCLB requires districts to set aside 20% of their yearly Title I allocated funds for school choice transportation and SES (Kim & Sunderman, 2005). Districts in improvement status are required to spend a minimum of 5% of the total set aside on SES. According to the U.S. Department of Education (2005e), "The per-child cost for Supplemental Educational Services is determined by dividing the district's Title I, Part A allocation by the number of children residing within the district aged 5 to 17 who are from families living below the poverty level." Costs paid out for SES are strictly related to provider fees. Title I funds are not permitted to be used for administrative costs or any other cost associated with implementing SES within a school or district. This controversial aspect of SES funding has resulted in claims that NCLB is an unfunded mandate.

Many educational leaders and state and local officials are critical of NCLB policies. In an effort to establish accountability, schools have drastically narrowed their curriculum, becoming intently focused on teaching reading and math (Fletcher, 2005). The system seems unresponsive to problems and views all the schools the same. Furthermore, fiscal considerations may discourage districts from promoting NCLB's choice options. The more that students pursue these options, the more districts will have to devote the mandatory 20% Title I budget set aside to SES programs rather than to programs already in existence with the likelihood that even 20% will not be adequate to cover the cost (Finn & Hess, 2004).

SES was established to be regulated by the state and local school districts in an effort to increase individual students' academic achievement through after-school tutoring for students in schools classified as needing improvement. As with any new program, monitoring and evaluation, data

---

collection and analysis are vital to the overall success of SES. Fortunately for SES providers, forerunners in SES implementation have identified “best practices” (Cohen, 2003; U.S. Department of Education, n.d.e).

### *Research Studies*

Very few studies have investigated the effects of NCLB on student learning outcomes. Due to the paucity of research in this area, each study shall be discussed in some detail. A study conducted by Bowe, Cain, Kingsbury, and McCall (2005) used the Growth Research Database from the Northwest Evaluation Association to compare student achievement and student growth on a common and reliable scale. The participants included hundreds of thousands of students in school districts across the country. The purpose of the study was to provide an initial view of the law and to identify trends. Findings from the studies indicated that state level tests tend to improve observed achievement and there is evidence that NCLB has improved student achievement since its adoption, although the effect is smaller than the testing effect. The measured growth in achievement may not necessarily be due to interventions on behalf of NCLB but may be attributed to the process of testing and retesting students and/or regular academic growth.

While NCLB has shown positive effects on student achievement and growth, there are two concerns raised by this study. The first is that at the current rate of change, schools will not be close to reaching the requirement of 100% proficiency by 2014. The second is that students in ethnic groups that have demonstrated achievement gaps in the past have had less growth under NCLB, and demonstrate less growth in comparison to European-American students with the same baseline score. NCLB was in its initial stages and it may have been too early in program implementation to identify the extent to which NCLB will influence educational change in the future.

Secretary Spellings (U.S. Department of Education, n.d.b) asserts NCLB has had the intended positive effect on students. The latest Nation’s Report Card (U. S. Department of Education, 2006a) shows steady growth and gains by students particularly among younger and minority students; overall fourth grade and eighth grade math scores increased as well as fourth grade reading scores. African-American and Hispanic fourth graders reached the highest reading and math scores for their groups than in any previous year, and African-Americans and Hispanic eighth graders reached the highest math scores for their groups than in any previous year. In both fourth and eighth grades, a higher percentage of white, African-American, Hispanic and Asian/Pacific Islander students performed at or above proficient than those in previous years. Although improvement has been measured, the results suggest middle and high school students may be an area of weakness.

While NCLB has garnered vast interest and many articles, there has been very little written about the SES portion of the initiative. Basically there are two areas of study pertain-

ing to SES. The first area investigates SES implementation and the second area looks at student achievement gains. One study presented key findings over a period of three years (David et al., 2006). Data were collected from a variety of sources including: a yearly survey of 1,300 district Title I administrators, a yearly survey of 739 principals, yearly case study visits, interviews of Title I administrators, and analyses of state accountability system components. From the study emerged five themes:

1. Small district schools were more likely to exit improvement status than large districts.
2. Participation in school choice remained at one percent and participation in Supplemental Educational Services increased from 7 to 19%.
3. An increased number of states provided technical assistance to schools in improvement status.
4. Strategies for school improvement remained similar across the three years nationally.
5. School poverty and district size were higher predictors of exiting improvement status than improvement strategies.

The biggest challenges districts who implemented SES faced included the lack of available providers (especially in small, rural districts), communication with parents, and assessing provider performance. In both 2002-03 and 2003-04, the number of eligible and participating students was substantially increased in urban and very large districts from 9,000 to 16,000 (on average) students. SES providers were primarily non-faith-based and non-online providers.

A study by Kim and Sunderman (2004) used 11 urban districts from a geographically, politically, and demographically diverse sample to provide a wide range of local contexts in which to examine the ability of districts to implement SES. The results of the study confirmed that SES was not widely used during the first year. The demand for services was low, primarily due to the inconvenience of services being offered outside of regular school hours and away from eligible students’ neighborhoods. The first year also documented tremendous administrative burdens faced by districts with no increased funding. Moreover, there was growing concern of the potential for SES to fragment Title I, seriously disrupting other school reform efforts by diverting resources away from the neediest students.

Another study (Anderson & Laguarda, 2005) used case studies conducted during 2003-04 school year and followed baseline data that were collected in the previous year. The study also conducted interviews in a purposive sample of six states and nine school districts, which were selected because they appeared to be relatively advanced in the process. Findings indicate that after two years, states, districts, schools, and providers were overcoming some of the initial trials of SES implementation. A noted area of improvement was establishing routines for reviewing applications and getting a list of providers out to districts sooner. District administrators continued to confront additional administra-

---

tive responsibilities and were in the process of developing systems that would streamline operations. Other challenges included moderate increases in participation, evaluating provider performance, improving communication with parents, managing administrative costs, and payment to providers when student attendance is irregular.

In a study conducted by the Chicago Public School Office of Research, Evaluation and Accountability (2005), baseline achievement levels of students who participated in the program were compared to eligible students who did not participate. The students with tutoring increased at national norms, while those students without tutoring had slightly fewer students at national norms. In addition to measuring student gains, gains between providers were examined. Students from one specific provider were shown to outperform students from other providers. Few researchers have examined the impact of Supplemental Educational Services on student achievement and school performance. There is no body of research that provides conclusive evidence documenting the effect of SES on learning outcomes for low-income and minority students. Does student participation in SES increase student learning outcomes in reading and math? The earliest implementation of SES was during the school year 2003-2004 and the impacts of which are in their infancy.

Few studies have been published that report the effects of Supplemental Educational Services on student outcomes, primarily because traditional statistical methods typically used in evaluations just won't work well with SES data. The purpose of this study is to implement an evaluation approach based on contemporary statistics that can effectively analyze SES data, a meta-analytic approach using effect size analysis. Our effective implementation of this evaluation approach may provide important information for future program evaluations and previously administered SES programs. Specifically, this study will add to the body of knowledge on educational reform to better understand the effectiveness of various SES programs. With any policy change, it is important to monitor the impact of that change and to evaluate the value of associated new programs. The primary focus of this study was to investigate the relationship between SES and achievement gains in a typical sample of SES students within Ohio and to identify strengths and weaknesses in provider practices that may be associated with student learning gains.

## Methodology: Analyses and Results

To determine if student achievement is enhanced by student participation in SES, the researcher examined three years of results from an ongoing statewide Evaluation Project and performed a meta-analytic analysis that estimated the overall impact of Supplemental Educational Services in predicting increased student achievement. As with all SES programs, the participants in the study were selected on the basis of need, and achievement performance data were collected at the start and at the end of SES programs.

The data groups for this study come from two public school districts located in the State of Ohio. SES programs

used in this study were implemented by local school district personnel. Due to the eligibility requirements for SES participation, it is assumed that individual SES samples are generally similar to SES populations nationwide, primarily comprised of high poverty, low achieving, minority students (U.S. Department of Education, 2005e). SES samples have been determined to be similar to SES populations nationwide (U.S. Department of Education, 2005e). Sufficient research and conclusive statistical evidence indicating SES populations are similar demographically in composition (U.S. Department of Education, 2005e). Students are eligible, whether or not they perform below the established level of proficiency, poorly on an assessment, or are part of a particular subgroup who attend Title I schools that have not met AYP for two consecutive years (Corwin & Wilhelm, 2006). National studies have characterized SES students as coming from low-income families, high poverty schools, and within the lower rankings for statewide assessments (U.S. Department of Education, 2004b, 2005e). Based on a theoretical framework and current research, it is assumed that the SES providers are generally comparable to other SES providers across the state and across the country.

The data were in the form of learning gains on academic assessments in reading and math. Individuals with incomplete records such as a pretest score and no posttest score were eliminated. A *t* test for dependent means was conducted to determine the statistical significance of the difference between the pre- and posttests for each group, see Table 1 for results. The data were analyzed with both descriptive and inferential statistics to determine irregularities in distribution. Means, standard deviation, skewness, and kurtosis were investigated and reported for critical variables.

Researchers are challenged regarding the use of statistical power and analysis when sample sizes are varied. According to Cohen (1992), "It is most useful to determine the *N* necessary to have a specified power for given (significance criterion) and ES" (p. 156). This problem frequently arises in ex post facto studies like the ones used in this study. In order to detect differences, as the sample sizes decreased in size the level of risk became greater.

An effect size functions as a descriptor to explicate the meaning of effect size. Cohen (1988, p. 25) described effect size as, "the average percentile standing of the average treated (or experimental) participant relative to the average untreated (or control) participant. At statisticians' disposal are two types of effect sizes; standardized difference and variance-accounted for effect sizes (Thompson, 2000). While an effect size correlation is calculated using original standard deviations, the Cohen's *d* uses a pooled standard deviation to calculate effect. There is reason to consider the impact of selecting effect size type. Dunlop, Cortina, Vaslow, & Burke (1996), Thompson (2000), and Rosenthal (1993) reported that pooled standard deviations tend to inflate the actual effect. Therefore, for this investigation, original standard deviations were used to compute effect size correlations.

In ex post facto research, it is necessary to determine the effect size (ES) in order to set the Type I error. The ES

was used to quantify the effectiveness of SES because it clarified the results and explained how well SES treatment worked (Cohen, 1988; Rosnow & Rosenthal, 1996). The ES allows results to be interpreted beyond statistical significance to practical implication and determine whether or not the results add to the general body of knowledge (Cohen, 1992). Statistical significance alone does not reveal the true size of the effect. The effect of SES on student achievement in math and reading was presumed to be large or noticeable to the observer.

In order to minimize the risk of Type I error, the investigator examined the necessary sample sizes suggested by Cohen (1992) for the specified level of significance and effect sizes. The suggested sample size for *t* test of dependent means with a large ES and level of significance .10 was 20 partici-

pants. The Ns for the groups with the level of significance of .10 were 11, 6, 46, and 17. From 2002 to 2004, SES participation increased from 7% to 19% (David et al., 2006). The rate of increase was reflected in this study's sample size. Groups F and H met the required sample sizes. In spite of insufficient sample sizes of groups A, B, D, E, and G, it was decided to proceed with the analysis. There were six reading groups (A through F) and two math groups (G and H), see Figure 1. The effect size (ES) for a large number of participants is assumed to be a more precise estimate of the population of the effect size based on a much smaller population. Because the sample sizes of the groups were so discrepant, each group was weighted by its number of participants.

Descriptive statistics and *t*-tests results for dependent mean comparisons, defined by pre and posttest mean differ-

Table 1  
*Pre- and Post-Test Distribution*

Sample	Subject	N	Measure	Mean	Std. dev.	Skewness	Kurtosis
A	Reading	11	Pretest	54.09	13.99	-.78	-.23
			Posttest	47.64	15.92	.96	.14
B		6	Pretest	353.33	7.00	.50	1.74*
			Posttest	382.00	30.35	-.06	1.74*
C		46	Pretest	225.43	77.04	-.25	-.42
			Posttest	271.96	75.14	.09	-.41
D		8	Pretest	4.75	1.03	-.39	-.49
			Posttest	5.50	1.06	.47	-.83
E		12	Pretest	11.83	8.10	.91	-.06
			Posttest	18.75	11.87	-.04	-1.46*
F		312	Pretest	2.66	1.69	.50	-.05
			Posttest	3.31	1.78	.91	-.55
G	Math	17	Pretest	7.47	6.85	2.48*	7.20*
			Posttest	12.00	9.51	1.45*	2.09*
H		119	Pretest	3.52	1.59	-.52	-.56
			Posttest	3.49	1.33	-.52	.15

Note: \*Indicates skewness or kurtosis outside of the acceptable 1.2 range.

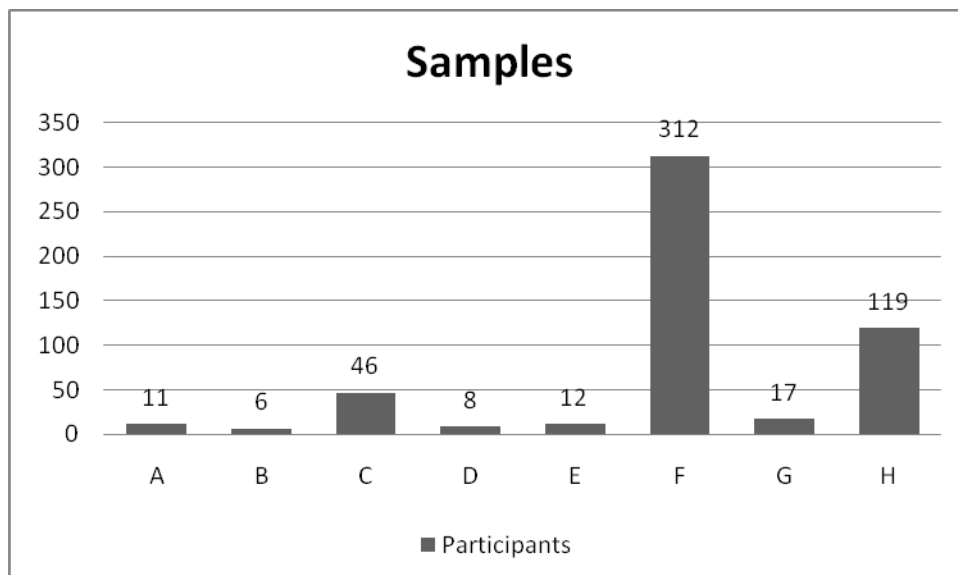


Figure 1. Sample distribution.

ences for each group, are reported in Table 2. Groups A and H do not show significant pre-post differences. Groups B, C, D, E, F, and G show significant differences,  $p < .10$ . The results of these analyses indicate significantly ( $p < .10$ ) greater achievement post scores on average for the SES groups.

In order to conduct an empirical evaluation and aggregate statistics, it is necessary to convert all of the summary statistics of the various studies into a common effect size. Using the statistics provided from the  $t$  tests of dependent means,  $t$  values were converted into Pearson Product Moment Correlations  $r$ . Where the original result indicated a negative effect, the posttest had a larger mean than the pretest; the  $r$  was treated as negative.

Confidence intervals for each sample's weighted effect size were computed and comparisons made to determine the stability of their effect size. Confidence intervals for every sample included zero, suggesting the distinct possibility of no effects. Because individual study sample sizes were relatively small, the level of significance was set at .05. Table 3 provides a summary index of the combined procedures of statistical significance of the  $t$ -test results.

Next, the effect size correlations were combined by averaging the raw Pearson correlation coefficients. The combined effect size correlation ( $r$ ) for reading = .71 and math = .01. At this point in the meta-analytic process, typically each  $r$  is transformed into a  $Z$  statistic using Fisher's  $r$  to  $Z$  transformation and then the  $Z$  scores are averaged and

transformed back into  $r$  values. There is some controversy involving this process. Fisher (1932) argued that  $Z$  scores somewhat overestimate  $r$  when sample sizes are small and  $r$  is large. A more conservative method (Hunter et al, 1982 and Rosenthal, 1984 & 1993) is to use a weighted average  $r$ . A weighted  $r$  is calculated by weighting each correlation by the number of subjects in that particular study. Taking into consideration Fisher's (1932) argument, the weighted  $r$  was used to combine effect sizes. The results of the combined weighted effect sizes ( $r^w$ ) were reading = .48 and math = -.04, results are shown in Table 3.

The heterogeneity of ES was examined to discern whether or not it was appropriate to synthesize the study results into one meta-analysis or if subsets should be considered. The most contributing factor in heterogeneity was sample size. To find the heterogeneity of the studies, the researcher calculated  $Q$  statistics and then distributed the scores in a chi-square. A chi-square was used to determine the degree of probability that the observed variance in ESs was the result of sample error alone. The criterion was set at .05. For reading the chi square was significant ( $\chi^2(5) = 12.67, p < .05$ ). The critical value for chi-square with  $df=5$  and  $p < .05$  is 11.07. The  $\chi^2$  is greater than the critical value indicating the variability across effect sizes does exceed what would be expected by just chance alone. This may be due to the degree of variation between sample sizes amongst studies. Sample F (with 312 subjects) contributes the most

Table 2  
*Paired Samples t Tests for Dependent Means*

Sample	<i>N</i>	Pre-test X	<i>sd</i>	Post-test X	<i>sd</i>	<i>t</i>	Dir. of effect	<i>p</i>	ES
A	11	54.0	13.9	47.6	15.9	1.17	-	.26	.21
B	6	353.3	7.0	382.0	30.3	2.27	+	.07*	.94
C	46	225.4	77.0	271.9	75.1	5.40	+	.00*	.39
D	8	4.7	1.0	5.5	1.0	4.58	+	.00*	.37
E	12	11.8	8.1	18.7	11.8	3.40	+	.00*	.32
F	312	2.6	1.6	3.3	1.7	19.73	+	.00*	.21
G	7	7.4	6.8	12.0	9.5	3.02	+	.00*	.27
H	119	3.5	1.5	3.4	1.3	0.81	-	.41	.04

Note. \*statistical significance at .10.

Table 3  
*Standard Values for the Samples*

Sample	Subject	<i>N</i>	<i>df</i>	<i>t</i>	ES <i>r</i>	wES <i>r</i> <sup>w</sup>	Lower Upper Confidence intervals
A	Reading	11	10	1.17	-0.35	-0.01	-0.69 to 0.69
B		6	5	2.27	0.71	0.01	-1.13 to 1.14
C		46	45	5.40	0.63	0.07	-0.29 to 0.31
D		8	7	4.58	0.87	0.02	-0.86 to 0.00
E		12	11	3.40	0.72	0.02	-0.65 to 0.67
F		312	311	19.73	0.75	0.59	-0.08 to 0.15
G	Math	17	16	3.02	0.60	0.08	-0.46 to 0.49
H		119	118	0.81	-0.07	-0.06	-0.19 to 0.18

to the heterogeneity of the studies. The studies are considerably heterogeneous, making it quite likely that a Type I error influenced the end result for the reading analysis. Due to the uneven sample sizes, a Type I error may indicate differences when in actuality there are none. A chi-square was calculated to determine the heterogeneity of math samples. The critical value for chi-square with  $df = 1$  and  $p < .05$  is 3.84. The  $\chi^2$  value for math is 0.15 and does not exceed the critical value indicating the distribution of effect sizes is homogenous and neither of the two groups needs to be removed.

Confidence intervals for ES and weighted-ES were computed and comparisons made to determine the stability of the effect size. The level of significance was set at .05. The confidence intervals for reading were 0.3791 to 0.5809 and do not include zero. Therefore, the probability that there is likely a true effect in reading exists. The confidence intervals for math were  $-0.2119$  to  $0.1319$  and do include zero. Thus, it is possible that the math effects may be a fluke, unlikely to replicate.

One shortcoming of a meta-analysis is that a researcher only has access to studies that have been published. Because an unknown amount of research remains unpublished and/or may end up locked in a file cabinet somewhere, this is known as the file drawer problem (Borg, Gall, & Gall, 2003). In order to combat this issue regarding the stability of results, one must calculate the Fail-safe N. The researcher computed the number of nonsignificant studies that would have to be hidden away or filed away in order to make the meta-analysis nonsignificant, see Table 4. If the Fail-safe N is larger than the determined number of studies, then the meta-analysis is considered valid. The Fail-safe N for reading was 812 studies and the critical number of nonsignificant studies 40. There would have to exist at least 812 nonsignificant studies to render the meta-analysis for reading insignificant because that number exceeds the critical number therefore no file drawer problem exists. The number of significant studies that would need to exist for math to be significant was 2 and the critical number of nonsignificant studies 20. The critical number exceeds the Fail-safe N and therefore the study suffers from a file drawer problem. It must be noted that none of the samples in this study have been published and in effect this study is the Fail-safe N. Table 4 provides a summary of the meta-analyses.

## Discussion

This study yielded relevant and applicable findings regarding student participation in SES, student achievement. A

meta-analysis was conducted to investigate the relationship between student achievement gains in reading and participation in SES. Six reading studies were used in the meta-analysis, with a total of 395 participants. A  $\chi^2$  was conducted to determine the heterogeneity of effect sizes. The results were  $\chi^2(5) = 12.67$ ,  $p < .05$  which is statistically significant. The samples were determined to be heterogeneous, most likely due to the difference in sample sizes. Sample F contributed most to the heterogeneity of the study. Although precautions were taken and the samples were weighted, the uneven distribution of participants and samples may contribute to rejecting the null when there is no true effect, a Type I error. The combined weighted effect size correlation = 0.48 and is considered a large effect. It is improbable the results are due to chance. It would have required an additional 812 nonsignificant studies to render the meta-analysis nonsignificant. The file drawer problem does not affect this study because only 40 additional nonsignificant studies are likely to exist. The CI range  $-0.38$  to  $0.58$  includes zero within its interval. Therefore, there is the possibility that the effect size is equal to zero, suggesting the possibility of no effects.

This study yielded relevant and applicable findings regarding student participation in SES, student achievement, and provider practices. A meta-analysis was conducted to investigate the student achievement gains in math experienced by students who participated in SES. Two studies were used in the meta-analysis, with a total of 136 participants. A  $\chi^2$  was conducted to determine the heterogeneity of effect sizes. The results were  $\chi^2(1) = 0.15$  that indicated the studies were homogeneous. The combined weighted effect size was  $r = 0.04$ , which Cohen (1992) considers insignificant. This may be in effect a Type II error and due to the paucity of studies to compare in the meta-analysis.

Participants in SES are predominantly from low-income families, high poverty schools, rank in the lower percentile on statewide assessments, and many are minority students (U.S. Department of Education, 2005). The students that are eligible for SES are a homogenous group most likely because our schools are failing these types of students. Schools have an obligation to teach all students and SES may be more suitable for this student demographic. The successes of SES programs should be examined and viable elements should be introduced into the regular classroom to assist this student demographic.

Educational practice is plagued with attempts to develop specific approaches to teaching that assume the approach will be effective in any setting and for all types of students

Table 4  
Meta Analysis Summary Table

Variable	No.	N	r	r <sup>1</sup>	wCI		Fail-safe N	Critical No.
					Lower Bound	Upper Bound		
Reading	6	395	.71	.48	0.3791 to 0.5809		812	40
Math	2	136	.01	.04	$-0.2119$ to $0.1319$		2	20

(Danielson, 1996). Investigating the variables associated with SES may provide data regarding the elements associated with academic gains. For example, if a lower student to teacher ratio in math is found to be a predictor of math learning gains, a school district may choose to incorporate breakout math classes for its students who are falling behind. In many districts, gifted children are sent to breakout classrooms for math or reading. These classes often have smaller teacher to student ratios. Offering the same type of program for remedial students may have the same effects as SES with the benefits of funding remaining within the school and the students' "school day" not being extended.

In conducting this study, certain observations have been made that could serve as the catalyst for future research. Data do not distinguish between regular academic growth and SES treatment. As it stands now, the numbers reflect regular academic growth with SES treatment. Without more accurate measures, we can only suggest what the effectiveness of NCLB policies might be, but we cannot be certain. Gains may be attributed to other factors such as new technology, professional development, or a new math series. Without proper data collection efforts in all states, the proficiency and effectiveness of the SES sanction remains undetermined.

Another area of focus for future research may be the differential effect between math and reading. Historically, there has been more extensive research done in the area of reading than in math. This contributes to the widespread and accepted research-based learning strategies and instructional practices in reading. This may be a contributing factor in the effectiveness of SES programming.

Are children better off having received SES? That question needs further analysis in order to justify the cost to public schools and the redirecting of Title I funding. Since the program is relatively new and there have been reported difficulties in implementation during the first 2 years, the researcher suggests further investigation regarding the impact of SES on both mathematics and reading scores in a longitudinal study. The study should examine learning gains in both reading and mathematics in an effort to determine the point of most dramatic affect on student achievement as well as long-term gain.

## References

- Anderson, L. M., & Laguarda, K. G. (2005). Case studies of supplemental services under the no child left behind act findings from 2003-2004. *U.S. Department of Education*. 1-85.
- Borg, W., Gall, J., & Gall, M. (2003). *Educational research: An introduction* (7<sup>th</sup> ed.). Boston: Pearson Education.
- Bowe, B., Cronin, J., Kingsbury, G., & McCall, M. (2005). *The impact of the NCLB act on student achievement and growth: 2005 edition*. Northwest Evaluation Association.
- Carr, F. (2001). A critique of school-to-work (stw) integration into traditional education subjects. *The Journal of Economics & Economic Education Research*. Arkansas State University: Whitney Press.
- Chicago Public Schools Office of Research, Evaluation & Accountability (2005). SES tutoring programs: An evaluation of the second year-part one of a two-part report. *Department of Evaluation and Data Analysis*.
- Clark-Carter, D. (1997). *Doing quantitative psychological research from design to report*. East Sussex, UK: Psychology Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Cohen, J. H. (2003). Supplemental education: Six essential components. *Principal*, 82(5), 34-37.
- Corcoran, T., & Goertz, M. (2005). The governance of public education. In S. Fuhrman & M. Lazerson (Eds.), *The public schools* (pp. 25-56). New York: Oxford University Press.
- Corwin, T., & Wilhelm, S. (2006). *Supplemental educational services web cast*. Retrieved June 25, 2006, from U.S. Department of Education.
- David, J., Laguarda, K., Lash, A., Lopez-Torkos, A., Padila, C., Shields, P., Skolnik, H., & Woodworth, K. (2006). *Title I accountability and school improvement from 2001 to 2004* (pp. 1-160). Washington, DC: U.S. Department of Education.
- Dunlop, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1, 170-177.
- Elmore, R. (2004). *School reform from the inside out, policy, practice, and performance*. Cambridge, MA: Harvard Education Press.
- Finn, C., Jr., & Hess, F. (2004). Inflating the life rafts of NCLB: Making public school choice and supplemental services work for students in troubled schools. *Phi Delta Kappan*, 86(1), 34-58.
- Fisher, R. A. (1932). *Statistical methods for research workers* (4<sup>th</sup> ed.). London: Oliver and Boyd.
- Fletcher, M. A. (2005). Federal mandate: Is the No Child Left Behind act making the grade? *Crisis*, 112(5), 16-19.
- Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (2007). *Supervision and instructional leadership: A developmental approach*. (7<sup>th</sup> ed.). Pearson Education.
- Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies*. Beverly Hills, CA: Sage.
- Kim, J., Orfield, G., Sunderman, G. L., & Tracey, C. A. (2004). Listening to teachers, classroom realities and no child left behind. *The Civil Rights Project by Harvard University*.
- Kim, J., & Sunderman, G. L. (2004). Increasing bureaucracy or increasing opportunities? School district experience with supplement educational services. *The Civil Rights Project by Harvard University*.



- Kim, J., & Sunderman, G. L. (2005). Measuring academic proficiency under the no child left behind act: Implications for educational equity. *Educational Researcher*, 34(8), 3-13.
- Kruse, S. (forthcoming). *The educational leader as decision maker: Processes, issues and considerations*.
- Kruse, S., Liang, X., & Beese, J. (2005). Evaluation report of supplemental education services. *External Evaluation Report to Ohio Department of Education*.
- Kruse, S., Liang, X., & Beese, J. (2006). Evaluation report of supplemental education services. *External Evaluation Report to Ohio Department of Education*.
- Kruse, S., Liang, X., & Widenbaugher, B. (2004). *Supplemental evaluation services external evaluation project report*. Akron, OH: The University of Akron.
- Ravitch, D. (1985). *The schools we deserve: Reflections on the educational crises of our times*. New York: Basic.
- Rosenthal, R. (1984). *Meta-analytic procedures for social research*. Beverly Hills, CA: Sage.
- Rosenthal, R. (1993). *Meta-analytic procedures for social research*. London: Sage.
- SES: The Basics*. Retrieved June 25, 2006, from [www.tutorforkids.org/basics.asp](http://www.tutorforkids.org/basics.asp)
- SES Pilot Programs*. Retrieved June 25, 2006, from [www.ed.gov](http://www.ed.gov)
- SES-Who Does What?* Retrieved June 25, 2006, [www.tutorsforkids.org](http://www.tutorsforkids.org)
- Smole, D. P. (2006a). Supplemental educational services for children from low-income families under ESEA title I-A. *CRS Report for Congress*. Congressional Research Service: The Library of Congress.
- Stullich, S., Eisner, E., McCrary, J., & Roney, C. (2006). National assessment of Title I: Interim report. Volume I: Implementation of Title I. *Institute of Education Sciences*, Washington, DC: National Center for Educational Evaluation and Regional Assistance.
- Sunderman, G. L., & Kim, J. (2004). Increasing bureaucracy or increasing opportunities? School district experience with supplemental educational services. *The Civil Rights Project Harvard University*.
- Sunderman, G. L., Kim, J., & Orfield, O. (2005). *NCLB meets school realities: Lessons from the field*. Thousand Oaks, CA: Corwin Press.
- Thompson, B. (2000). *A suggested revision to the forthcoming 5<sup>th</sup> edition of the APA publication manual*. Retrieved August 30, 2008, from <http://www.coe.tamu.edu/~bthompson/apaeffect.htm>
- U.S. Department of Education (2002). *No Child Left Behind: Supplemental educational services, non-regulatory guidance*. Washington, DC: Author.
- U.S. Department of Education (2003). *No Child Left Behind*. Washington, DC: ED Pubs, Publications Center.
- U.S. Department of Education (2004a). *Supplemental educational services policy*. Washington, DC: ED Pubs, Publications Center.
- U.S. Department of Education (2004b). *Creating strong supplemental educational services programs* (pp. 1-63). Washington, DC: ED Pubs, Publications Center.
- U.S. Department of Education (2005a). *Commission Staff Research Report, Growth models: An examination within the context of NCLB. Commission on No Child Left Behind*. The Aspen Institute. Retrieved September 5, 2006, from [www.nclbcommission.org](http://www.nclbcommission.org)
- U.S. Department of Education (2005b). *Nation's report card shows continued progress*. Retrieved June 27, 2006, from [www.ed.gov](http://www.ed.gov).
- U.S. Department of Education (2005c). *No Child Left Behind: What parents need to know*. Washington, DC: Department of Education.
- U.S. Department of Education (2005d). *No Child Left Behind: A roadmap for state implementation*. Washington, DC: Department of Education.
- U.S. Department of Education (2005e). *No Child Left Behind: Supplemental educational services non-regulatory guidance*. Washington, DC: Department of Education.
- U.S. Department of Education (2006a). *The facts about... Measuring progress*. Retrieved 2006 from [www.ed.gov/print/nclb/accuntability/ayp/testing.html](http://www.ed.gov/print/nclb/accuntability/ayp/testing.html)
- U.S. Department of Education (2006b). *Digest of education statistics*. Washington, DC: National Center for Educational Statistics.
- U.S. Department of Education (n.d.a). Budget proposal advances NCLB, introduces competitiveness initiative. *The Achiever*, 5(3).
- U.S. Department of Education (n.d.b). *Key policy letters signed by the education secretary or deputy secretary*. Retrieved June 25, 2006, from [www.ed.gov/print/elsec/guid/secletter/060515.html](http://www.ed.gov/print/elsec/guid/secletter/060515.html)
- U.S. Department of Education (n.d.c). *Letter to chief state school officers, supplemental educational services policy: LEA conditions on providers*. Retrieved October 2006, from <http://www.ed.gov/policy/elsec/guid/stateletters/choice/ses082604.html>
- U.S. Department of Education (n.d.d). *Office of elementary and secondary education and office of innovation and improvement, supplemental educational services (SES) policy*. Retrieved May, 2006, from <http://www.ed.gov/policy/elsec/guid/stateletters/choice/ses051006.html>
- U.S. Department of Education (n.d.e). *Policy and program studies service report highlights: Early implementation of supplemental educational services under the No Child Left Behind act: year one report*. Retrieved 2006 from [www.ed.gov/rschstat/eval/disadv/supplementalyear1](http://www.ed.gov/rschstat/eval/disadv/supplementalyear1)
- Walberg, H. J. (2001). Achievement in American schools. In T. Moe (Ed.), *A primer on American schools* (pp. 43-98). Stanford, CA: Hoover Institution Press.
- Wilhelm, S., Rees, N., & Corwin, T. (2006). Supplemental educational services web cast.

---

# *The Impact of the Collective Efficacy of a School Community on Individual Professional Development Outcomes*

Cynthia H. Geer  
Xavier University  
Julie Q. Morrison  
University of Cincinnati

## *Abstract*

*In the past decade, leading contributors in the areas of educational change and staff development have acknowledged that successes attained with individual aspects of professional development can be stifled, halted, or essentially canceled by seemingly immutable factors in the organization's culture. The purpose of this study was to explore the impact of the collective efficacy of science teachers within a school community on the individual outcomes for science teachers participating in a professional development program. The findings suggest that teachers participating in the program from schools with high levels of collective efficacy demonstrated gains in science teaching self-efficacy that were greater than the gains shown by teachers in schools with low collective efficacy. This preliminary finding has potentially serious implications regarding the impact of a school's culture and collective beliefs on the professional development outcomes of its members.*

On-going professional development is essential for educators, not from an acknowledgement of deficiencies, but instead from recognition of education as a dynamic, ever advancing, professional field (Guskey, 2000). Consistent in the research on educational outcomes is the finding that notable improvements in student learning almost never take place in the absence of professional development (Guskey, 2000). The *No Child Left Behind Act* in 2001 increased the demands for professional development by mandating that school districts provide "highly qualified" teachers for every student. Specifically, this legislation states, "Each state education agency (SEA) must develop a plan to ensure that all teachers are 'highly qualified' no later than the end of the 2005-06 school year... In general, a 'highly qualified teacher' is a teacher with full certification, a bachelor's degree, and demonstrated competence in subject knowledge and teaching skills" (U.S. Department of Education, 2002, p. 57). In order to assist SEA's with meeting this goal, Title II funds are allocated for the purpose of preparing, training, and recruiting high quality teachers. The way these funds are used is at the discretion of the districts and schools in an effort to ensure that each is able to meet their individual needs, provided that the funds are used for "scientifically based professional development interventions" (U.S. Department of Education, 2002, p. 57).

### *School Factors Influencing Professional Development*

The influence of a school's culture and organization on the outcome of the professional development of its teacher members has received minimal attention to date. In the past decade, leading contributors in the areas of educational change and staff development have acknowledged that suc-

cesses attained with individual aspects of professional development can be stifled, halted, or essentially canceled by seemingly immutable factors in the organization's culture (Fullan, 1993; Sparks, 1996; Sparks & Hirsh, 1997). For professional development programs to maximize their success, careful consideration must be given to organizational elements that include: organizational policies (aligned with organization's mission), resources, protection from intrusions, openness to experimentation and alleviation of fears, collegial support, principal's leadership and support, higher-level administrators' leadership and support, recognition of success, provision of time (Guskey, 2000). Despite these discussions of the importance of attending to school factors in support of individual teachers' professional development, there has been no research to date examining the degree to which a school's culture influences professional development outcomes.

### *Theoretical Framework*

The theoretical foundation of this study rests with the construct of self-efficacy, as conceptualized by Albert Bandura's social cognitive theory. Self-efficacy is defined as a belief in one's own abilities to perform an action or activity necessary to achieve a goal or task (Bandura, 1997). Beliefs about one's own abilities to accomplish specific tasks are powerful predictors of behavior. Self-efficacy beliefs influence choices, effort, and persistence in the face of adversity (Pajares, 1997).

As it relates to teachers, teacher self-efficacy is the belief that one can bring about desired outcomes of student engagement and learning, even among those students who may be difficult or unmotivated (Tschannen-Moran, & Hoy, 2001). Teacher self-efficacy has been linked to teachers'

---

responsibility for student achievement (Guskey, 1982, 1988) and greater persistence in working with struggling students (Gibson & Dembo, 1984). Teacher self-efficacy is predictive to a willingness to implement innovative teaching strategies and improve methods of instruction (Allinder, 1994; Guskey, 1984, 1988; Smylie, 1988; Stein & Wang, 1988). Of particular interest to this study, the school environment has been shown to influence teachers' self-efficacy beliefs (Basom & Frase, 2004; Haney, Wan, Keil, & Zoffel, 2007; Tschannen-Moran, 2001).

A teacher's self-efficacy has been found to play a critical role in their ability to impact student achievement (Ashton, 1985; Ashton & Webb, 1986). With particular respect to math and science education, several studies have shown a positive association between teacher self-efficacy and elementary students' achievement in science (Cannon & Scharmann, 1996), student motivation in mathematics for students transitioning to junior high school (Midgley, Feldlaufer, & Eccles, 1989), and achievement in computer technology (Ross, Hogaboam-Gray, & Hannay, 2001). Given the positive student outcomes associated with higher levels of self-efficacy, an increase in teachers' self-efficacy appears to be reasonable objective of a professional development program that also sought to increase content knowledge and pedagogical skills (Geer & Morrison, 2007; Haney et al., 2007; Watson, 2006).

### *Collective Efficacy*

The construct of collective efficacy has emerged from the research on teacher self-efficacy. Rather than focusing on the individual perceptions of teachers regarding their own capacity, the unit of analysis is the group (or school faculty) for collective efficacy (Goddard, Hoy, & Woolfolk Hoy, 2000). Collective efficacy is defined as the "perceptions of group members about a faculty's conjoint capacity to successfully educate students" (Goddard, 2002, p. 98). The assessment of collective efficacy involves the aggregation of individual teachers' judgments of the capacity of the members of their school community as a whole to educate the students with whom they provide instruction.

The collective efficacy model used in this study is based on the model of teacher efficacy conceptualized by Tschannen-Moran and Hoy (2001). According to this model of teacher efficacy, two factors contribute to teachers' judgments of self-efficacy: (a) their perceptions of personal competence and (b) the demands of the task. These two factors are represented in the collective efficacy model as Group-Teaching Competence (GC) and Task Analysis (TA). GC consists of judgments about the capabilities that faculty members bring to a given teaching situation, whereas TA includes beliefs about the demands of teaching given the level of support provided by the students' home and community (Goddard, 2002).

### *Purpose of the Study*

The purpose of this study was to explore the impact of the collective efficacy of science teachers within a school

community on the individual outcomes for teachers participating in a professional development program for science educators. This study expands the research literature on professional development for teachers by quantifying the much-discussed influence of the school culture on professional development outcomes, and providing a theoretical foundation to conceptualize this influence. Three research questions were addressed in this study:

1. To what degree did participation in the professional development program increase teachers' self-efficacy in teaching science, compared to their self-efficacy in teaching science prior to the program?
2. To what degree did teachers who participated voluntarily in the professional development program differ in their self-efficacy in teaching science, compared to the science teachers in their school who did not participate directly in the program?
3. To what degree did the collective efficacy of the school influence the science teachers' gains in self-efficacy?

## Method

### *Professional Development Program Description*

Seventy-seven Catholic elementary schools located in a Midwestern city were invited to submit a proposal for participation in a five-year professional development program for improving teaching strategies and content knowledge in science and mathematics. Funding for the professional development program was granted by a private foundation in collaboration with a local Catholic university. Teams from each school consisted of at least one mathematics teacher, at least one science teacher, and a school principal. The professional development program included monthly sessions on pedagogy during the academic year. The monthly workshops examined topics that included, but were not limited to, constructivist learning theory; the learning cycle; national, state, and district content standards; pedagogical practices (e.g., assessment, inquiry-based instructional techniques, and the use of technology to enhance instruction). The summer science program, co-taught by a scientist and a science educator, was designed to increase content knowledge.

The 21 schools selected for participation in the professional development program were all Catholic elementary schools serving students in Grades 1-8. These urban, suburban, and rural schools varied by way of the socio-economic status and ethnic backgrounds of the students they served (See Appendix). Each school sent two cohorts of math and science teachers. Teachers were provided graduate credit, release time for professional development during the school day, and a modest stipend for their participation in the program.

Twenty-four science teachers from the 21 elementary schools comprised the first cohort to participate in the program. Among these teachers there were 23 females and 1 male. The prior teaching experience of the science teachers

ranged from 0-39 years, with an average of 14 years of teaching experience (an average of 9 years teaching science). Two science teachers discontinued their participation as a result of their school withdrawing from the professional development program (one of the schools was closed). Four additional science teachers left the program, one of which was replaced by an alternate. Consequently, 18 Cohort 1 science teachers completed the professional development program over 24 months. The program completion rate for Cohort 1 was 75.0%. A similar completion rate was obtained for Cohort 1 math teachers; however, this study focused exclusively on science teacher participants.

Fourteen of the original 21 schools continued their participation in the professional development program by sending a second cohort of math and science teachers. The decrease in number of schools in the Cohort 2 was due to school closings resulting from district restructuring, teacher transfers or career changes, and the challenges of recruiting additional teachers interested in the professional development program. The program duration was shortened from 24 months for Cohort 1 teachers to 18 months for Cohort 2 teachers to maximize the cost-effectiveness of the program delivery; however, the professional development program remained the same fundamentally. The teachers in Cohort 1 received additional pedagogical topics on informal science education, planning a science fair and a science family night, and literature in science teaching and learning.

Fourteen science teachers from 14 remaining elementary schools participated as members of the second cohort of the program. Among these teachers there were 11 female teachers and 3 male teachers. The prior teaching experience of the science teachers ranged from 0-35 years, with an average of 12 years of teaching experience (an average of 9 years teaching science). Three science teachers withdrew from the 18-month program for personal reasons. The program completion rate for Cohort 2 was 78.6%. A similar completion rate was obtained for Cohort 2 math teachers; however, this study focused exclusively on science teacher participants.

### *Participants*

*Cohort 1 and Cohort 2 Science Teachers.* The participants in this study were recruited as part of their schools' involvement in (and completion of) the professional development program. The final sample was comprised of 16 Cohort 1 teachers (a response rate of 88.9%) and 12 Cohort 2 teachers (a response rate of 85.7%).

*Non-Participant Science Teachers.* In addition, 37 teachers taught science in a school participating in the professional development program, but were not directly involved in the program as a Cohort 1 or Cohort 2 teacher. The majority of these teachers were female (90.8%) and White (90.8%). The teachers reported having 17.7 years of experience teaching, on average, with an average of 14.2 years of experience teaching science. The response rate for science teachers who were not themselves program participants was 89.0%. This group

represented a convenience sample as they were recruited by virtue of being colleagues of the program participants employed in the same school building.

### *Measures and Analyses*

*Science Teaching Efficacy Belief Instrument (STEBI).* The STEBI, developed by Riggs and Enochs (1990), was used to measure the teachers' judgments of their self-efficacy in teaching science. The STEBI consisted of 25 items using a 5-point Likert rating scale, where 5 = "Strongly Agree," 4 = "Agree," 3 = "Uncertain," 2 = "Disagree," and 1 = "Strongly Disagree." Twelve items were recoded to reverse the values associated with the responses and the item values were summed. Adequate reliability was demonstrated for the use of the STEBI in this study ( $\alpha = .88$ ).

Changes in teachers' self-efficacy were analyzed by comparing the teachers' self-ratings at the beginning of the professional development program with the teachers' self-ratings at the end of the program. Paired-sample *t*-tests were conducted to determine the degree to which the changes in the mean self-ratings were statistically significant. Effect sizes were calculated to determine the strength of the mean change, using the following formula: the difference between the mean of the post-measure and the mean of the pre-measure, divided by the standard deviation of the pre-measure. In addition, independent-sample *t*-tests were conducted to compare post-test only differences in self-efficacy in teaching science between the professional development participants and science teachers in their school who did not participate directly in the program.

*Collective Efficacy Scale—long version (CES)* (Goddard, Hoy, & Woolfolk Hoy, 2000). The CES was used to assess the collective efficacy of the science teachers in the participating schools. The scale provides for the measurement of the two factors of the collective efficacy model previously described, group-teaching competence (GC) and task analysis (TA). The CES is comprised of 21 items using a 6-point Likert rating scale, where 6 = "Strongly Agree" and 1 = "Strongly Disagree." Ten items were recoded to reverse the values associated with the responses and the item values were summed. Adequate reliability was demonstrated for the use of the CES in this study ( $\alpha = .90$ ).

### *Design & Procedures*

This study employed a within-groups pre-test post-test quasi-experimental design to compare gains in science teaching efficacy among teachers directly participating in the program. Teachers directly involved in the program (Cohort 1 & 2 teachers) completed the STEBI as a pre-test measure at the beginning of their participation in the program.

A post-test only comparison group design was used to compare ratings on the STEBI for teachers in the professional development program with their non-participant counterparts. The program teachers and their non-participant teaching counterparts completed the STEBI as a post-test measure and

the CES at the time Cohort 2 teachers completed the professional development program. At the same point in time that the Cohort 2 teachers completed the STEBI and the CES, Cohort 1 teachers and any remaining non-program science teachers were also administered the STEBI and the CES, for the purpose of this study. Table 1 provides an overview of the measurement occasions by group membership for the

Table 1  
*Overview of the Measurement Occasions by Group (Cohort 1 & 2 vs. Non-Participants)*

	N	Pre-Test Measure	Program Participation	Post-Test Measure
<b>STEBI</b>				
Cohort 1 & 2	16	X	Yes	X
Non-Participants	37		No	X
<b>CES</b>				
Cohort 1 & 2	28		Yes	X
Non-Participants	37		No	X

X = Measurement occasion

two instruments used in this study.

The CES measure for the school was obtained by aggregating the CES total score across all of the science teachers at that school who completed the CES. Between-group comparisons were made on the STEBI for science teachers in high-CES schools (scoring 67<sup>th</sup> percentile or higher relative to other schools in the sample) versus teachers in low-CES schools (scoring 33<sup>rd</sup> percentile or lower relative to other schools in the sample).

## Results

*Research Question 1. To what degree did participation in the professional development program increase teachers' self-efficacy in teaching science, compared to their self-efficacy in teaching science prior to the program?*

Sixteen science teachers of the 28 science teachers who participated in the professional development program as part of Cohort 1 and 2 completed both a pre-test measure and a post-test measure of the STEBI. The response rate was 57.1%, reflecting the challenges of collecting data longitudinally. These teachers increased their total score on the STEBI from a pre-test mean of 59.63 ( $SD = 8.47$ ) to a post-test mean of 97.81 ( $SD = 10.59$ ). The difference in the pre-test and post-test scores was statistically significant at the .05 level ( $t = 10.03, p = .00, df = 15$ ). Across both cohorts an effect size of 4.51 was obtained, indicating an extremely strong effect.

*Research Question 2. To what degree did teachers who participated voluntarily in the professional development program differ in their self-efficacy in teaching science, compared to the science teachers in their school who did not participate directly in the program?*

The mean total score on the STEBI for science teachers participating in the professional development program in

Cohort 1 and 2 ( $M = 97.58, SD = 11.55$ ) was not statistically significantly different than the mean total score for science teachers from the same schools who did not participate directly in the professional development program ( $M = 96.92, SD = 10.41$ ). Given that this finding is based on measures administered after the completion of the STEBI, it is likely that the teaching self-efficacy for the science teachers in Cohort 1 and 2 was significantly lower than their non-participant peers prior to the professional development program. It is also plausible that the non-participant peers demonstrated gains in science teaching self-efficacy during the same period of time as an indirect effect of their schools' participation in the professional development program or as a result of the Hawthorne Effect.

*Research Question 3. To what degree did the collective efficacy of the school influence the science teachers' gains in self-efficacy?*

Sixty-five science teachers from 14 elementary schools completed the CES. Of these participants, 28 teachers were Cohort 1 and 2 participants in the professional development program and 37 teachers taught science in a school participating in the program, but were not directly involved in the professional development program. The number of teachers who responded for each individual school ranged from two to six, depending on the size of the school faculty. CES total scores were aggregated by school and the mean CES ranged from 79.33-105.75, for an overall mean of 96.81 ( $SD = 12.53$ ). The mean GC factor score across 14 schools was 64.97 ( $SD = 4.07, range = 57.33-70.00$ ) and the mean TA factor score was 31.11 ( $SD = 4.76, range = 22.67-36.25$ ).

Among the four schools with mean CES total scores at or below the 33<sup>rd</sup> percentile for the group of schools, the teachers participating in the professional development program from those schools demonstrated a pre-test to post-test mean gain on the STEBI of 14.33 ( $SD = 14.22$ ). In contrast, among the four schools with mean CES total scores at or above the 67<sup>th</sup> percentile for the group of schools, the teachers participating in the program from those schools demonstrated a pre-test to post-test mean gain on the STEBI of 42.5 ( $SD = 10.75$ ). The difference in gain on the STEBI between teachers in high-CES schools and low-CES schools was statistically significant at the .05 level ( $t = -3.008, p = .03, df = 5$ ).

In interpreting the results to this third research question, it is important to recognize that the relationship between individual science teaching self-efficacy among the science teachers of a given school and the collective efficacy of the school as judged by the same science teachers is most likely reciprocal, and not uni-directional, in nature. Individual teaching self-efficacy may impact judgments of collective efficacy just as collective efficacy may influence individual teaching self-efficacy. Among the teachers in this study, the relationship between scores on the STEBI and the CES was strong ( $r = .78, p = .00$ ), suggesting that the two instruments measure attitudes that are conceptually interdependent, yet distinct. In this study, the collective efficacy of the school, as judged by the science teachers, was a far better predictor

---

of the science achievement of the students at sixth grade ( $r = .83, p = .00$ ) than was the STEBI as a predictor of sixth grade science achievement ( $r = .54, p = .00$ ), as measured by the Terra Nova standardized test. Plausible rival explanations for the findings in this study, the limitations of study, and recommendations for future research will be discussed in the next section.

## Discussion

As a nation, we spend \$5 to \$12 billion annually on the professional development of educators (Miles, Odden, Fermanich, & Archibald, 2004). The financial costs of teacher professional development are further compounded by the loss to students of instructional time provided by their teacher attending a professional development seminar (Siegel & Yates, 2007). These considerations, along with a growing body of evidence that professional development can influence teacher's knowledge and practice (Garet, Porter, Desimone, Birman, & Yoon, 2001; Supovitz & Turner, 2000) provide a strong basis for asserting that only the most productive professional development should be implemented. Furthermore, the degree to which the school organization supports the pursuit of professional development among its members is critical to maximizing the effects of the professional development.

The purpose of this study was to explore the impact of the collective efficacy of science teachers within a school community on the individual outcomes for teachers participating in a professional development program for science educators. Science teachers who participated directly in a professional development program designed to increase science teachers' content knowledge and pedagogical skills demonstrated gains in their self-efficacy in teaching science that were both statistically significant and practically significant. These teachers, who were self-selected to participate in a professional development program, did not differ significantly from their non-participant peers in their ratings of science teaching self-efficacy upon completion of the program. Coupled with the previous finding of gains in science teaching self-efficacy for the participants, it is possible that the teaching self-efficacy for the science teachers in the professional development program was significantly lower than their non-participant peers prior to the program. This might suggest a selection bias among the teachers participating in the program, favoring participants who initially perceived their self-efficacy for teaching science as low were more inclined to seek professional development. A second possible explanation is that the teachers who did not participate directly in the professional development program benefited *indirectly* from the experience of having their colleague within the same school share new knowledge and skills, as was encouraged in the professional development program. It is also plausible that the non-participant peers demonstrated gains in science teaching self-efficacy during the same period of time as a result of the Hawthorne Effect.

Future research should employ a between-group pre-test post-test comparison group design to determine if the groups are truly equivalent prior to the program implementation and assess gains in both groups throughout the professional development program. Ultimately, it would also be interesting to attempt to quantify the indirect influence of a professional development program on the non-participant peers who teach in the same subject area and to differentiate this influence from the Hawthorne Effect.

Findings from the third research question suggest that teachers participating in the professional development program from high-CES schools demonstrated gains in science teaching self-efficacy that were statistically significantly greater than the gains shown by teachers in low-CES schools. Although this is an interesting preliminary finding, much more rigorous research is needed to better understand the nature of the relationship between individual teachers' judgments of their teaching self-efficacy and teachers' perceptions of the collective efficacy of the school building, particularly as these variables impact and are impacted by professional development. In this study, the relationship between science teaching self-efficacy and collective efficacy was strong, but not perfect, suggesting that the two instruments measure attitudes that conceptually interdependent, yet distinct. What remains unknown is the degree to which these two potentially powerful predictors of student achievement influence each other. Does the collective efficacy of the school influence the teaching self-efficacy of the individual teacher? Does the teaching self-efficacy of individual faculty members influence the collective efficacy of the school. Mostly likely this relationship between teaching self-efficacy and collective efficacy is reciprocal, and not uni-directional. A better understanding of this relationship appears to be fertile ground for those interested in maximizing the benefits of professional development for teachers. If new knowledge and skills acquired through intensive professional development are not encouraged when the teacher returns to apply them in the school, the benefits of professional development will not be fully realized.

The distinction between high- and low-CES deserves further exploration. Does the high achievement of the students produce a sense of high collective efficacy? Does low collective efficacy diminish the quality of instruction provide to students, thus negatively impacting achievement? The explanation lies in social cognitive theory's concept of reciprocal determinism. Accordingly, personal factors, behavior, and environmental influences all operate interactively as determinants of one another (Bandura, 1986). When the unit of analysis is the individual, personal factors (which include beliefs, such as self-efficacy) bear on an individual's behavior (teaching practices), which in turn influence personal factor. Both personal factors and behavior determine in part the environment (student learning, collegial relationships), which in turn impacts personal factors and behavior. Therefore, individual teachers do not simply react passively to school situations; they actively create their own

school environments and act to change them. When the unit of analysis is the group, the collective beliefs of the school faculty, instructional and leadership behaviors, and school environment/organizational aspects and outcomes all operate interactively as determinants of one another. Future research should examine, on a larger scale, the relationship between CES and achievement with special attention to the outliers (i.e., schools with high CES and low achievement, schools with low CES and high achievement). Attending to both collective efficacy and student achievement might be the key to maximizing the impact of professional development for educators. Given that the stakes are high for our students, professional development must target, in the words of Coburn (2003), “deep and consequential change in classroom practice – change that goes beyond surface structures or procedures (such as changes in materials, classroom organization, or the addition of specific activities) to alter teachers’ beliefs, norms of social interaction, and pedagogical principles as enacted in the curriculum” (p. 5). Collective efficacy holds tremendous promise for improving student outcomes for all learners.

## References

- Allinder, R. M. (1994). The relationships between efficacy and instructional practices of special education teachers and consultants. *Teacher Education and Special Education, 17*, 86-95.
- Ashton, P. T. (1985). Motivation and the teacher’s sense of efficacy. In C. Ames & R. Ames (Eds.), *Research on motivation in education: Vol. 2. The classroom milieu* (pp. 141-174). Orlando, FL: Academic Press.
- Ashton, P. T., & Webb, R. B. (1986). *Making a difference: Teachers’ sense of efficacy and student achievement*. New York: Longman.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.
- Basom, M., & Frase, L. (2004). Creating optimal work environments: Exploring teacher flow experiences. *Mentoring & Tutoring: Partnership in Learning, 12*(2), 241-258.
- Cannon, J., & Scharmann, L. C. (1996). Influence of a cooperative early field experience on preservice elementary teachers’ science self-efficacy. *Science Education, 80*, 419-436.
- Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32*(6), 3-12.
- Fullan, M. G. (1993). *Change forces: Probing the depths of educational reform*. Bristol, PA: Falmer.
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915-945.
- Geer, C. H., & Morrison, J. Q. (2007). *Increasing teacher self-efficacy and decreasing the achievement gap: Professional development to improve science learning in Catholic elementary schools*. Paper presentation at the Association for Science Teacher Education International Conference, Clearwater Beach, FL.
- Gibson, S., & Dembo, M. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology, 76*, 569-582.
- Guskey, T. (1982). Differences in teachers’ perceptions of personal control of positive versus negative student learning outcomes. *Contemporary Educational Psychology, 7*, 70-80.
- Guskey, T. (1984). The influence of change in instructional effectiveness upon the affective characteristics of teachers. *American Educational Research Journal, 21*, 245-259.
- Guskey, T. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education, 4*(1), 63-69.
- Guskey, T. R. (2000). *Evaluating Professional Development*. Thousand Oaks, CA: Corwin Press.
- Goddard, R. (2002). A theoretical and empirical analysis of the measurement of collective efficacy: The development of a short form. *Educational and Psychological Measurement, 62*(1), 97-110.
- Goddard, R. D., Hoy, W. K., & Woolfolk Hoy, A. (2000). Collective teacher efficacy: Its meaning, measure, and effect on student achievement. *American Education Research Journal, 37*(2), 479-507.
- Haney, J., Wang, J., Keil, C., & Zoffel, J. (2007). Enhancing teachers’ beliefs and practices through problem-based learning focused on pertinent issues of environmental health science. *Journal of Environmental Education, 38*(4), 25-33.
- Midgley, C., Feldlaufer, H., & Eccles, J. (1989). Change in teacher efficacy and student self- and task-related beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology, 81*, 247-258.
- Miles, K. H., Odden, A., Fermanich, M., & Archibald, S. (2004). Inside the black box of school district spending on professional development: Lessons from comparing five urban districts. *Journal of Education Finance, 30*(1), 1-26.
- Pajares, F. (1997). Current directions in self-efficacy research. In M. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 1-49). Greenwich, CT: JAI Press.
- Riggs, I., & Enochs, L. (1990). Towards the development of an elementary teacher’s science teaching efficacy belief instrument. *Science Education, 74*, 625-637.

- Ross, J., Hogaboam-Gray, A., & Hannay, L. (2001). Effects of teacher efficacy on computer skills and computer cognitions of Canadian students in Grades K-3. *The Elementary School Journal, 102*(2), 141-156.
- Siegel, W., & Yates, C. (2007). *Improving the evaluation of professional development presentations using retrospective pre-testing of existing knowledge and self-efficacy*. National Council of Professors of Educational Administration. Retrieved January 21, 2008, from <https://cnx.org/content/m14362/latest>
- Smylie, M. A. (1988). The enhancement function of staff development: Organizational and psychological antecedents to individual teacher change. *American Educational Research Journal, 25*, 1-30.
- Sparks, D. (1996, February). Viewing reform from a systems perspective. *The Developer*, pp. 2, 6.
- Sparks, D., & Hirsh, S. (1997). *A new vision for staff development*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Stein, M. K., & Wang, M. C. (1988). Teacher development and school improvement: The process of teacher change. *Teaching and Teacher Education, 4*, 171-187.
- Supovitz, J.A., & Turner, H. M. (2000). The effects of professional development on science teaching practices and classroom culture. *Journal of Research in Science Teaching, 37*(2), 963-980.
- Tschannen-Moran, M. (2001). The effects of a state-wide conflict management initiative in schools. *American Secondary Education, 29*(2), 2-33.
- Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783-805.
- U.S. Department of Education, Office of Elementary and Secondary Education (2002). *No Child Left Behind, a desktop reference*. Washington, DC.
- Watson, G. (2006). Technology professional development: Long-term effects on teacher self-efficacy. *Journal of Technology & Teacher Education, 14*(1), 151-165.

## Appendix

### Demographic Characteristics of the Participating Schools

#### Race/Ethnicity

School	Student Enrollment	African American	American Indian or Native Alaskan	Asian or Pacific Islander	Hispanic	Multi-racial	White	Economically Disadvantaged	Region
SC01	140	0.0%	0.0%	0.0%	0.7%	0.7%	98.6%	11.4%	Urban
SC02	594	7.7%	0.5%	1.2%	0.3%	1.9%	88.4%	4.5%	Suburban
SC03	165	0.0%	0.0%	0.6%	1.2%	0.0%	98.2%	1.8%	Rural
SC04	171	5.8%	0.6%	1.2%	2.9%	7.0%	82.5%	35.7%	Urban
SC05	220	96.8%	0.0%	0.0%	0.0%	3.2%	0.0%	85.9%	Urban
SC06	416	13.9%	0.0%	2.4%	0.2%	5.0%	78.4%	3.1%	Suburban
SC07	1030	0.2%	0.0%	1.1%	0.0%	0.1%	98.6%	7.0%	Suburban
SC08	164	7.3%	0.0%	3.0%	0.0%	5.5%	84.1%	13.4%	Suburban
SC09	249	3.6%	0.0%	4.4%	0.4%	6.0%	85.5%	57.0%	Urban
SC10	192	0.0%	0.0%	0.5%	0.5%	4.7%	94.3%	15.6%	Rural
SC11	505	2.2%	0.0%	0.8%	1.2%	0.6%	95.2%	5.3%	Urban
SC12	373	10.7%	0.0%	0.0%	1.1%	7.2%	81.0%	9.7%	Urban
SC13	155	11.6%	0.0%	5.8%	2.6%	7.1%	72.9%	27.1%	Suburban
SC14	99	91.9%	0.0%	0.0%	1.0%	2.0%	5.1%	56.6%	Urban
SC15	192	0.0%	0.0%	0.0%	0.5%	0.5%	99.0%	24.5%	Suburban
SC16	126	33.3%	0.0%	0.0%	0.0%	7.9%	58.7%	38.9%	Urban
SC17	131	19.1%	.08%	0.0%	0.0%	0.8%	79.4%	90.8%	Urban
SC18	711	0.0%	0.0%	0.1%	0.4%	1.4%	98.0%	6.3%	Suburban
SC19	855	0.0%	0.0%	0.0%	0.0%	0.0%	100%	5.6%	Suburban
SC20	364	3.3%	0.0%	0.3%	0.3%	5.2%	90.9%	6.6%	Suburban
SC21	246	2.0%	0.0%	0.0%	0.4%	1.2%	96.3%	23.6%	Suburban



---

## *Index of Articles: 2007–2008*

- AYP Accountability Policy and Assessment Theory Conflicts*, Volume 21, No. 4, Fall 2008  
Randall S. Davies, Brigham Young University
- Bouncing Back after Bullying: The Resiliency of Female Victims of Relational Aggression*, Volume 21, No. 2, Spring 2008  
Laura R. Hammel, Cleveland State University and Ursuline College
- Choosing to Serve? An Exploration of Student Self-Selection of Service Learning Projects*, Volume 21, No. 2, Spring 2008  
Paula S. Weber, St. Cloud State University  
Kenneth R. Schneider, St. Cloud State University  
James E. Weber, St. Cloud State University
- The Effect of Supplemental Educational Services on Student Learning Outcomes*, Volume 21, No. 4, Fall 2008  
Jane Beese, University of Akron
- An Exploratory Survey of Participants in Urban and Suburban Teacher Academy Programs*, Volume 21, No. 4, Fall 2008  
Leah Wasburn-Moses, Miami University  
Molly Kelly-Elliott, Miami University
- Frequency of Principal Turnover in Ohio's Elementary Schools*, Volume 21, No. 2, Spring 2008  
Michelle Chaplin Partlow, Temple University  
Carolyn S. Ridenour, University of Dayton
- The Heart of the Scholarship of Teaching and Learning—Lifelong Mentoring*, Volume 21, No. 1, Winter 2008  
Sharon Valente, Ashland University
- The Impact of the Collective Efficacy of a School Community on Individual Professional Development Outcomes*, Volume 21, No. 4, Fall 2008  
Cynthia H. Geer, Xavier University  
Julie Q. Morrison, University of Cincinnati
- MWERA 2007 Highlights*, Volume 21, No. 1, Winter 2008  
Dimitar M. Dimitrov, George Mason University
- The Road to Getting Published*, Volume 21, No. 1, Winter 2008  
Patricia B. Elmore, Southern Illinois University Carbondale  
Darren M. James, Southern Illinois University Carbondale
- The Roles and Responsibilities of MWERA Participants: A Personal Perspective*, Volume 21, No. 3, Summer 2008  
Katrina M. Daytner, Western Illinois University
- Standards in Conducting and Publishing Research in Education*, Volume 21, No. 1, Winter 2008  
Bruce Thompson, Texas A&M University and Baylor College of Medicine
- Teacher Efficacy and Career Indecision among Pre-Service Teachers: A Model of Direct and Indirect Effects*, Volume 21, No. 2, Spring 2008  
Mary Beth Slone, University of West Georgia  
Mary D. Hancock, University of West Georgia
- Thirty Years with MWERA and a View of the Future*, Volume 21, No. 1, Winter 2008  
Jean Pierce, Northern Illinois University
- Understanding School Board Members' Perceptions of Superintendents' Leader Behaviors*, Volume 21, No. 4, Fall 2008  
John V. Richard, The University of Akron  
Sharon D. Kruse, The University of Akron
- What Constitutes Good Educational Research? A Consideration of Ethics, Methods and Theory*, Volume 21, No. 1, Winter 2008  
Kelly D. Bradley, University of Kentucky  
Kenneth D. Royal, University of Kentucky  
Jessica D. Cunningham, University of Kentucky  
Jennifer A. Weber, University of Kentucky  
Jennifer A. Eli, University of Kentucky

---

## Index of Authors: 2007–2008

- Beese, Jane**, University of Akron  
*The Effect of Supplemental Educational Services on Student Learning Outcomes*, Volume 21, No. 4, Fall 2008
- Bradley, Kelly D.**, University of Kentucky  
*What Constitutes Good Educational Research? A Consideration of Ethics, Methods and Theory*, Volume 21, No. 1, Winter 2008
- Cunningham, Jessica D.**, University of Kentucky  
*What Constitutes Good Educational Research? A Consideration of Ethics, Methods and Theory*, Volume 21, No. 1, Winter 2008
- Davies, Randall S.**, Brigham Young University  
*AYP Accountability Policy and Assessment Theory Conflicts*, Volume 21, No. 4, Fall 2008
- Daytner, Katrina M.**, Western Illinois University  
*The Roles and Responsibilities of MWERA Participants: A Personal Perspective*, Volume 21, No. 3, Summer 2008
- Dimitrov, Dimiter M.**, George Mason University  
*MWERA 2007 Highlights*, Volume 21, No. 1, Winter 2008
- Eli, Jennifer A.**, University of Kentucky  
*What Constitutes Good Educational Research? A Consideration of Ethics, Methods and Theory*, Volume 21, No. 1, Winter 2008
- Elmore, Patricia B.**, Southern Illinois University Carbondale  
*The Road to Getting Published*, Volume 21, No. 1, Winter 2008
- Geer, Cynthia H.**, Xavier University  
*The Impact of the Collective Efficacy of a School Community on Individual Professional Development Outcomes*, Volume 21, No. 4, Fall 2008
- Hammel, Laura R.**, Cleveland State University and Ursuline College  
*Bouncing Back after Bullying: The Resiliency of Female Victims of Relational Aggression*, Volume 21, No. 2, Spring 2008
- Hancock, Mary D.**, University of West Georgia  
*Teacher Efficacy and Career Indecision among Pre-Service Teachers: A Model of Direct and Indirect Effects*, Volume 21, No. 2, Spring 2008
- James, Darren M.**, Southern Illinois University Carbondale  
*The Road to Getting Published*, Volume 21, No. 1, Winter 2008
- Kelly-Elliott, Molly**, Miami University  
*An Exploratory Survey of Participants in Urban and Suburban Teacher Academy Programs*, Volume 21, No. 4, Fall 2008
- Kruse, Sharon D.**, The University of Akron  
*Understanding School Board Members' Perceptions of Superintendents' Leader Behaviors*, Volume 21, No. 4, Fall 2008
- Morrison, Julie Q.**, University of Cincinnati  
*The Impact of the Collective Efficacy of a School Community on Individual Professional Development Outcomes*, Volume 21, No. 4, Fall 2008
- Partlow, Michelle Chaplin**, Temple University  
*Frequency of Principal Turnover in Ohio's Elementary Schools*, Volume 21, No. 2, Spring 2008
- Pierce, Jean**, Northern Illinois University  
*Thirty Years with MWERA and a View of the Future*, Volume 21, No. 1, Winter 2008
- Richard, John V.**, The University of Akron  
*Understanding School Board Members' Perceptions of Superintendents' Leader Behaviors*, Volume 21, No. 4, Fall 2008
- Ridenour, Carolyn S.**, University of Dayton  
*Frequency of Principal Turnover in Ohio's Elementary Schools*, Volume 21, No. 2, Spring 2008
- Royal, Kenneth D.**, University of Kentucky  
*What Constitutes Good Educational Research? A Consideration of Ethics, Methods and Theory*, Volume 21, No. 1, Winter 2008
- Schneider, Kenneth R.**, St. Cloud State University  
*Choosing to Serve? An Exploration of Student Self-Selection of Service Learning Projects*, Volume 21, No. 2, Spring 2008
- Slone, Mary Beth**, University of West Georgia  
*Teacher Efficacy and Career Indecision among Pre-Service Teachers: A Model of Direct and Indirect Effects*, Volume 21, No. 2, Spring 2008
- Thompson, Bruce**, Texas A&M University and Baylor College of Medicine  
*Standards in Conducting and Publishing Research in Education*, Volume 21, No. 1, Winter 2008
- Valente, Sharon**, Ashland University  
*The Heart of the Scholarship of Teaching and Learning—Lifelong Mentoring*, Volume 21, No. 1, Winter 2008
- Wasburn-Moses, Leah**, Miami University  
*An Exploratory Survey of Participants in Urban and Suburban Teacher Academy Programs*, Volume 21, No. 4, Fall 2008
- Weber, Jennifer A.**, University of Kentucky  
*What Constitutes Good Educational Research? A Consideration of Ethics, Methods and Theory*, Volume 21, No. 1, Winter 2008
- Weber, James E.**, St. Cloud State University  
*Choosing to Serve? An Exploration of Student Self-Selection of Service Learning Projects*, Volume 21, No. 2, Spring 2008
- Weber, Paula S.**, St. Cloud State University  
*Choosing to Serve? An Exploration of Student Self-Selection of Service Learning Projects*, Volume 21, No. 2, Spring 2008

# ***Mid-Western Educational Researcher Reviewers for 2007-2008***

*We want to thank the individuals listed below who served as reviewers for the past year*

Abdel Kader, Shereen	Harris Stowe State University	Newman, Isadore	University of Akron
Allen, Kimberly	University of Missouri	Newman, Mary	College of DuPage
Atauhene, Francis	Western Illinois	Ngudgratoke, Sungworn	University of Michigan
Backus, Angela	Kent State	Parker, Abby	Ohio State University
Baptiste, Lennise	Kent State	Paul, Kelli	Indiana University
Barrow, Lloyd	University of Missouri	Pearson, Bernadina	George Mason University
Batagiannis, Stella	Indiana University- Purdue	Pierce, Jean	Northern Illinois University
Bouck, Emily	Purdue University	Pokay, Patrick	Eastern Michigan University
Boyer, Patrick	University of Missouri	Ramanathan, Hema	University of West Georgia
Brown, Kathleen	University of Missouri	Rausch, John	John Carrol University
Bruening, Paige	Ohio State University	Royal, Kenneth	University of Kentucky
Cody, Thomas	Western Illinois	Sheehan, Heather	Bowling Green State University
Corfman, Steven	Antioch College	Silverman, Sarah	Ohio State University
Damore, Sharon	DePaul	Sigler, Ellen	Indiana University
Davis, Stephanie	Wright State	Slutsky, Ruslan	University of Toledo
Edlefson, Carla	Western Illinois	Soboleski, Penny	University of Findlay
Falk-Ross, Fran	Pace University	Sonmez, Serpil	Florida State University
Fereday, Linda	University of Kentucky	Smith, LaTisha	Harris Stowe State
Floyd, Eva	University of Kentucky	Staple, Christopher	University of Kentucky
Franco, Suzanne	Wright State	Still, Kristin	Cleveland State
Fordham, Nancy	Bowling Green State University	Stuckey, Angeline	Northern Illinois University
Gerrity, Kevin	Ball State	Stucky-Mickell, Tracey	Northern Illinois University
Gunn, Kelly	University of Cincinnati	Todaro, Stacey	Northern Illinois University
Headley, Stacey Anne	Ohio State University	Vander Schee, C	Northern Illinois University
Henning, Mary Beth	Northern Illinois University	Valente, Sharon	Ashland University
Hofmann, Richard	Miami University	Vonderwell, Selma	Cleveland University
Jones, Nathan	Michigan State University	Wang, Aimin	Miami University
Kantner, Joanne	Kishwaukee College	Was, Christopher	Kent State University
Kessinger, Thomas	Xavier University	Weber, Jennifer	University of Kentucky
Kim, Yonghee	Ohio State University	Wendel, Paul	Kent State University
Leatherman, Jane	Indiana University - Purdue	Wiggins, Kathryn	DePaul University
Lee, Hea-Jin	Ohio State University	Walker, David	Northern Illinois University
McNeely, Sharon	Northeastern University	Williams, Jane	Ohio State University
Mertler, Craig	Bowling Green State University	Williams, Nicole	Ohio State University
Miller, Angela	Ball State University	Xia, Mingzhu	University of Tennessee
		Zhao, Yu	Pennsylvania State University
		Zimmerman, Judy	Bowling Green State University



The Ohio State University  
2500 Kenny Road  
Columbus, Ohio 43210

Nonprofit Org.  
U.S. Postage  
PAID  
Columbus, Ohio  
Permit No. 711