10-9-2012

A Quasi-Experimental Test of the Elementary School Success Profile Model of Assessment and Prevention

Natasha K. Bowen
University of North Carolina at Chapel Hill

Aaron M. Thompson
University of Missouri

Joelle D. Powers
Boise State University

Journal of the Society for Social Work and Research
Volume 3, Issue 3, 178-196

October 9, 2012
ISSN 1948-822X DOI:10.5243/jsswr.2012.12

A Quasi-Experimental Test of the Elementary School Success Profile Model of Assessment and Prevention

Natasha K. Bowen
University of North Carolina at Chapel Hill

Aaron M. Thompson
University of Missouri

Joelle D. Powers
Boise State University

The Elementary School Success Profile Model of Assessment and Prevention (ESSP MAP) is an assessment and intervention strategy designed to improve student academic performance and behavior. The current analysis uses a quasi-experimental design to examine the relationship between a 3-year implementation of the ESSP MAP and aggregate academic outcomes. Students in one 3rd grade cohort (2007-2008) from 4 schools in 1 district received the intervention as they progressed from 3rd to 5th grade. Longitudinal hierarchical linear modeling with cross-classified effects of schools (N = 10) and cohorts (N = 11) was used to compare trajectories of reading proficiency percentages for the targeted group overall and its demographic subgroups with the trajectories of analogous groups of students across schools and time in the district. Findings suggest that the ESSP MAP was associated with greater growth in reading proficiency rates for Black and White students.

Key words: elementary school, achievement gap, universal prevention, hierarchical linear modeling, cross-classified random effects

The academic proficiency of American elementary school students as measured by federal standardized tests continues to be unacceptably low in spite of more than a decade of federal, state, and local efforts to boost achievement. Academic proficiency is mastery of the knowledge and skills in an academic area necessary for doing grade level work (National Center for Education Statistics [NCES], 2011a), and is the minimum desirable level of competency for students. In 2011, only 34% of American fourth graders were at or above proficiency in reading, and 40% were at or above proficiency in math (NCES, 2011b). Although low rates of proficiency are cause for concern, discrepancies among subgroups are even more alarming. For example, in 2011, only 16% of Black and 19% of Latino fourth-grade students were proficient in reading as compared with 44% of White fourth-grade students (NCES, 2011b). Gaps associated with family income were also pronounced: in 2011, 18% of students eligible for free or reduced-price lunches at school were proficient in reading as compared with 48% of noneligible students (NCES, 2011b).

The federal government has attempted to promote better student academic outcomes through the No Child Left Behind Act (NCLB, 2001) and the reauthorized Individuals with Disabilities Education Act (IDEA, 2004). Among other things, NCLB mandated that states regularly evaluate school performance using standardized tests, and that results be disaggregated by racial/ethnic, economic, and disability status. IDEA defines the categories of disability for which students are entitled to special educational services. Special education, or exceptional children's services, refers to entitlements funded with federal dollars to help schools support students with serious behavioral, academic, and emotional disorders. Both legislative agendas seek to improve performance outcomes overall while reducing the behavioral and academic gaps among student subgroups by using regular assessments and scientifically based educational interventions (Sugai & Horner, 2009). To accomplish these ends, both pieces of legislation encourage the use of tiered-response frameworks to address student needs and avoid overreliance on special education services for students who are responsive to less-intensive interventions. Features of tiered-response systems are described in more depth below, but the essentials include (a) school-wide instruction provided to all students using scientifically based academic and social programs and practices, (b) ongoing monitoring of student responses to school-wide instructional programs and practices, and (c) additional supports for students who struggle to meet adequate academic and social functioning as measured by ongoing data (Fuchs & Fuchs, 2005; Malecki & Demaray, 2007; Sugai & Horner, 2009). One goal of tiered-response systems is to provide additional support services to

students who need such services without prematurely or inappropriately referring students for expensive and lengthy evaluations for exceptional children's services (Fuchs, Mock, Morgan, & Young, 2003; Gresham, 2005).

In reality, many school systems are ill-equipped to implement the tiered prevention and intervention framework intended by federal legislative mandates (Alonzo, Tindal, & Robinson, 2008; Christ & Hintze, 2007; Kovaleski, 2007), which may help explain the persistence of low academic performance. To effectively implement a tiered framework, local educators need a thorough understanding of the process as well as access to appropriate resources for data collection, decision making, and intervention implementation. Two specific barriers to effective implementation of tiered frameworks in schools are (a) the failure to use social environmental assessments to guide the first tier of prevention efforts, and (b) a lack of support in the form of materials, training, personnel, and funding for data collection, decision making, and the implementation of interventions.

The current study examined the effects of a practice model that reduces barriers associated with the implementation of tiered prevention frameworks. The Elementary School Success Profile Model of Assessment and Practice (ESSP MAP) was implemented with one cohort of third-grade students in 4 of 10 schools in one district. The outcome of interest was growth in the percentage of students who demonstrated proficiency on their state standardized reading tests from third to fifth grade. The outcome was compared across intervention and nonintervention cohorts and schools. End-of-grade tests are the standardized academic tests administered to third through eighth grade students at the end of each academic year in North Carolina to meet federal accountability mandates. We were particularly interested in the effects of the ESSP MAP on increases in the percentages of proficient students among Blacks and Latinos and students from low-income families. These three groups have consistently had lower proficiency rates in the district (North Carolina Department of Public Instruction [NCDPI], 2011), and the United States as a whole (NCES, 2011b). Overall improvements in proficiency rates of the targeted district were not possible unless growth was seen in the three subgroups. Due to space considerations and heterogeneity of the disability subgroup, the current study does not encompass an examination of outcomes for students with disabilities.

Background and Theory

Response to Intervention (RtI) and Positive Behavioral Interventions and Support (PBIS) are two tiered prevention frameworks that have gained prominence in school systems. Although the two frameworks are similar, the focus of RtI is often on academics and the focus of PBIS is often on behavior. Tiered models include early screening, use of evidence-supported interventions, and ongoing collection of data to monitor student progress and determine the need for more intensive services or entitlements (Batsche et al., 2005; Fuchs et al., 2003; Gresham, 2005). In theory, the process starts with early assessment of malleable social and environmental risk and protective factors known to affect the social and academic success of students. Educators then select universal (or Tier 1) strategies to reduce identified risk factors or sustain protective factors to prevent the development of academic or behavioral problems. Prior research has shown that faithful implementation of effective universal programs supports the success of approximately 80% of students (Frey, Lingo, & Nelson, 2008; Malecki & Demaray, 2007; Wilson & Lipsey, 2008).

While universal strategies are in place, ongoing performance data (e.g., referrals to the office for misbehavior, ongoing academic assessment information; Burns & Gibbons, 2008) are collected on all students. Educators then systematically determine whether students who fail to respond to universal strategies require services that are more intensive. These students may be considered at-risk for developing serious academic or social problems but are expected to succeed with secondary (or Tier 2) prevention efforts (Frey et al., 2008; Malecki & Demaray, 2007). Prior research has suggested a subset of approximately 10% to 15% of students who will likely require more intensive Tier 2 interventions (Wilson & Lipsey, 2008). An even smaller subset of students (approximately 5%) might fail to respond adequately to either Tier 1 or Tier 2 interventions. These students might be referred for intensive individual interventions (Tier 3). Fewer than this 5% of students may appropriately be identified as needing exceptional children's services (Sugai & Horner, 2009).

Researchers across multiple disciplines have identified individual and social-environmental influences acting as risk and protective factors in the development and functioning of children (e.g., Bronfenbrenner, 1979; Fraser, Kirby, & Smokowski, 2005; Sameroff, 2000). For example, negative characteristics of and experiences in the neighborhood, family, school, and peer system are associated with low academic performance and behavior problems (Case & Katz, 1991; Malecki & Demaray, 2007; Malecki & Elliot, 2002; Pritchard & Wilson, 2003; Richman, Bowen, & Woolley, 2004; Spicker, Southern, & Davis, 1987). The knowledge base regarding risk and protective factors indicates that low income is a risk factor for school difficulties, as is minority

racial/ethnic status and largely due to the association of minority status in the United States with low income (Richman et al., 2004).

Consistent with empirical evidence of the impact of social environmental factors on functioning, most evidence-based interventions for youth target the social environment, such as parenting practices, school climate, or peer-group characteristics (e.g., Henggeler & Borduin, 1990; Olweus, Limber, & Mihalic, 1999; Webster-Stratton, 1997) and the individual characteristics that promote successful interactions between the individual and the environment, such as social skills or self-regulation (Lochman, Coie, Underwood, & Terry, 1993). However, contrary to the multidisciplinary body of research supporting early assessment of the social environment and the long history of tiered prevention (Gordon, 1983), the initial universal screening step used in most schools to inform tiered approaches typically relies on academic or behavioral indicators (Kovaleski, 2007; Kratochwill, Clements, & Kalymon, 2007). Illustrating this point, the National Center on Response to Intervention (2010) has listed 48 "screening" tools, each of which is a measure of current reading or math performance. Such measures are important for ongoing assessment in a tiered-response framework; however, these measures are indicators only of the nature and magnitude of performance problems. Measures of reading or math performance do not inform educational professionals of the malleable environmental risk factors that contribute to and help explain poor academic performance.

In summary, identifying ecological threats to performance is key to understanding and reducing the threats. A screening focus on problems instead of causes of problems is a major barrier to the proper implementation of successful tiered supports for students who are at-risk of academic failure. Without universal preventive supports at Tier 1, the identification of students needing supports at Tiers 2 and 3 is unlikely to occur as intended; causes of problems will remain unaddressed and more students than necessary will need targeted services.

A second barrier contributing to the difficulties local school professionals encounter in properly implementing a tiered-response prevention framework involves a lack of tangible support for newly mandated activities. Although NCLB (2001) and IDEA (2004) endorsed the use of RtI and PBIS frameworks to improve student outcomes, reduce the achievement gap, and decrease the number of students referred for entitlement services, the policies did not provide a funding mechanism to pay for training school staff in data collection methods, using data for decision making, or choosing and implementing appropriate interventions. Nor do the policies pay for the acquisition of valid and reliable measures or evidence-based intervention programs (e.g., copyrighted materials, supplies, additional staffing) necessary to achieve legislative goals (Luo, 2008; Scott & Martinek, 2006; Shen & Cooley, 2008). Without training and access to appropriate assessments and interventions, schools are unlikely to develop the capacity to carry out the steps necessary to achieve the goals outlined in NCLB and IDEA.

ESSP MAP

The ESSP MAP (Bowen, 2011; Bowen, Bowen, & Woolley, 2004; Bowen & Powers, 2011; Powers, Bowen, & Bowen, 2011; Webber, Rizo, & Bowen, 2011; Wegmann, Thompson, & Bowen, 2011) is an approach that addresses the previously mentioned barriers to effectively implementing a tiered-prevention framework in schools. Although the ESSP MAP could be readily integrated into all three tiers of an RtI or PBIS model in a school, the ESSP MAP was used in the current study independent of RtI or PBIS and with a predominantly universal (Tier 1) purpose. The ESSP MAP includes an online social environmental assessment administered to parents, students, and teachers, and a set of online resources and guidelines for school staff. The online Elementary School Success Profile (ESSP) assessment tool (Bowen, 2011; Bowen et al., 2004; Webber et al., 2011; Wegmann et al., 2011) collects data from the three sources about risk and protective factors related to the neighborhood (parent and child perceptions), school (parent and child perceptions), peer system (child's perception), family (parent and child perceptions), parent educational behavior (parent and teacher perceptions), health and well-being of the student (child perceptions), home and school social behavior (parent and teacher perceptions), and school performance (teacher perceptions). The ESSP is the elementary version of the School Success Profile, which was developed for middle and high-school students in the 1990s (Bowen, Richman, & Bowen, 2002; Bowen, Rose, & Bowen, 2005; Richman et al., 2004). The ESSP generates school-, group-, and individual-level profiles of social environmental experiences and self-perceptions of students in Grade 3 through Grade 5. The current study focuses on the ESSP group-level data as a guide for choosing Tier 1 strategies, but individual-level data are also useful for guiding Tier 3 intervention efforts.

Online materials include guidelines for interpreting the profiles, templates for identifying target areas, guidelines for writing achievable intervention goals, an extensive online database (Powers et al., 2011) of evidence-based and promising school-based strategies

to address identified risk factors publicly available at www.schoolsuccessonline.com and templates for planning and monitoring the implementation fidelity of selected strategies.

Similar to the tiered intervention approaches promoted in federal legislation, the ESSP MAP is not a specific intervention or combination of interventions. Rather, the ESSP MAP is a framework that organizes resources for school personnel to collect and use relevant data to categorize and organize supports that are responsive to the needs of their students. Because the specific needs of students vary across time, school, student group, and grade level, we argue that data-driven, tailored responses to student needs are appropriate in schools, even though the variability of interventions that occur in any one year across schools or within one school across time presents an evaluation challenge. For the current study, we hypothesized that providing schools with ESSP data, training on how to use the data and how to select appropriate interventions, and discretionary funding to support the implementation of universal interventions would lead to improvements in school-level academic outcomes.

Method

Design

The study used a multiple nonequivalent comparison group, longitudinal, quasi-experimental design. Three schools with the highest percentages of free-lunch program participants (see Table 1) and lowest percentages of students passing state standardized tests were purposively chosen to receive the intervention by administrators in an above-average school district in North Carolina. District administrators believed ESSP MAP resources would be most effectively applied to the district's lowest performing elementary schools. In Year 2 of the study, a newly built fourth school was added to the intervention condition because some students who had been enrolled in the original study site schools were transferred to the new school during their fourth-grade year. Similar to the original three schools, when the new school was added to the intervention group, the school had a relatively low reading proficiency rate (84.9%) and a relatively high percentage of free-lunch program participants (24%) as compared with the nonintervention schools in the district. The three original schools implemented the ESSP MAP with the 2007-2008 cohort of third graders and continued the intervention with those third graders as the students moved through the fourth and fifth grades. The fourth school joined the study in the 2008-2009 school year and conducted the intervention with the same student cohort starting in the students' fourth-grade year. Although the schools receiving the ESSP MAP had the highest rates of lunch program participation in their district, those rates (about 33% or less) were lower than the rates of lunch program participation of many schools in the state. Therefore, the external validity of the study might be limited in relation to high-poverty schools and districts.

The outcome used in the study was the trajectory of end-of-grade reading proficiency percentages associated with each cross-classified cohort and school, that is, groups defined by both cohort membership and school. Each trajectory with complete data comprised three data points: the percentage of cohort members at a school who were proficient in reading on their Grade 3 end-of-grade tests, the percentage of members of the same cohort in the same school the following year who were proficient in reading on their Grade 4 end-of-grade tests, and the percentage of members of the same cohort in the same school who were proficient in reading on their Grade 5 end-of-grade tests. Aggregate cohort scores at each grade level were based on the students at the school at the time of testing. Four trajectories were associated with the ESSP MAP: one for the 2007-2008 cohort of third graders for each of the four schools that received the intervention.

We used data from 10 other cohorts and six other elementary schools in the district to create comparison trajectories. Cohort 1 contained students who were third graders in the 2000-2001 school year in all schools in the district. We chose the 2000-2001 school year as the starting point because it was the year in which NCLB (2001) was passed, with its emphasis on tiered-response models to improve school level accountability for the performance of student subgroups. This choice was a logical starting point because we wanted to compare the effects of ESSP MAP with the patterns of performance since NCLB was put in place. With 10 schools and 11 cohorts, 110 school-level trajectories were possible. However, due to the construction of two schools in the district after the 2000-2001 year, eight potential trajectories were missing (i.e., one comparison school had no data for the first cohort and the fourth ESSP MAP school had no data for the first seven cohorts). Therefore, we had data on 102 trajectories of third-to-fifth grade reading proficiency percentages: four intervention trajectories and 98 comparison trajectories. The design is illustrated in Table 2.

Table 1
Known Selection Bias: Poverty and Reading Proficiency Rates of Third to Fifth Graders in 2006-2007 (Year Before Intervention), by School

	% Free or Reduced Price Lunch	% All Students Proficient
ESSP MAP School 1	33.0	89.8
ESSP MAP School 2	27.0	88.8
ESSP MAP School 3	31.7	88.7
ESSP MAP School 4	NA[a]	NA[a]
Comparison School 1	20.4	95.0
Comparison School 2	26.1	92.0
Comparison School 3	17.9	95.0
Comparison School 4	17.8	93.3
Comparison School 5	19.6	91.8
Comparison School 6	11.0	95.0

Note. [a]Target school #4 opened in the 2008-09 academic year

The 98 comparison trajectories included three subgroups: (a) trajectories for all cohorts before and after the target cohort in the four intervention schools (n = 33), (b) trajectories for students in the target cohort (2007-2008 third graders) but not in intervention schools (n = 6), and (c) trajectories for non-ESSP cohorts and non-ESSP schools (n = 59). Due to sample size limitations, scores for the three comparison subgroups could not be compared separately to ESSP MAP trajectories; the three subgroups were combined into one comparison group big enough to help compensate for the small size of the treatment group.

The use of multiple "non-equivalent, comparison groups" has been recommended by Shadish, Cook, and Campbell (2002, p. 159). Others have suggested that across- and within-cohort designs produce results with moderate to strong internal and external validity, control for history or contextual changes, and manage differences in outcomes associated with student socio-demographics (Bryk, Raudenbush, & Ponisciak 2004; Ponisciak & Bryk, 2005). The three comparison subgroups mitigated different threats to the internal validity of the study as described by Shadish and colleagues (2002) and others. Subgroup (a) reduced the selection threat that other cohorts in the intervention schools had similar trajectories of improvement prior to the intervention, or that change during the intervention period represented regression to the mean. Subgroup (b) reduced history and instrumentation threats, that is, the possibility that outcomes were due to the intervention group's experiences with changes in the school system, the standardized testing tool used that year, or any other outside influence unique to the years of the study. Comparison subgroup (c) reduced the maturation threat to internal validity, that is, that the pattern of improvement over time was typical for the district as a whole or for third through fifth graders in general. Because all students in both conditions completed multiple standardized tests during their third to fifth grade careers, the design also controlled for testing effects.

Another threat to internal validity was managed statistically. Because the statewide cutoff point for reading proficiency was increased during the first year of the study, we included a statistical control for that year's cohort of students. Although the recalibration should have affected the school-level proficiency scores of all schools in similar ways, the control reduced the chance that any treatment effects found were due to differential effects of the recalibration on the intercepts or slopes of school-level trajectories in low- and high-performing schools. The control variable is described more fully in the Measures section.

Table 2

Outcome Observation and Intervention Schedule of Multiple Nonequivalent Comparison Cohort Design

	Cohort's 3rd Grade	Cohort's 4th Grade	Cohort's 5th Grade
Cohort 1: 2000-2001 3rd Graders	Spring 2001	Spring 2002	Spring 2003
ESSP Schools (3 trajectories)	O_1	O_2	O_3
Non-ESSP Schools (5 trajectories)	O_1	O_2	O_3
Cohort 2: 2001-2002 3rd Graders	Spring 2002	Spring 2003	Spring 2004
ESSP Schools (3 trajectories)	O_1	O_2	O_3
Non-ESSP Schools (6 trajectories)	O_1	O_2	O_3
Cohort 3: 2002-2003 3rd Graders	Spring 2003	Spring 2004	Spring 2005
ESSP Schools (3 trajectories)	O_1	O_2	O_3
Non-ESSP Schools (6 trajectories)	O_1	O_2	O_3
Cohort 4: 2003-2004 3rd Graders	Spring 2004	Spring 2005	Spring 2006
ESSP Schools (3 trajectories)	O_1	O_2	O_3
Non-ESSP Schools (6 trajectories)	O_1	O_2	O_3
Cohort 5: 2004-2005 3rd Graders	Spring 2005	Spring 2006	Spring 2007
ESSP Schools (3 trajectories)	O_1	O_2	O_3
Non ESSP Schools (6 trajectories)	O_1	O_2	O_3
Cohort 6: 2005-2006 3rd Graders	Spring 2006	Spring 2007	Spring 2008
ESSP Schools (3 trajectories)	O_1	O_2	O_3
Non-ESSP Schools (6 trajectories)	O_1	O_2	O_3
Cohort 7: 2006-2007 3rd Graders	Spring 2007	Spring 2008	Spring 2009
ESSP Schools (3 trajectories)	O_1	O_2	O_3
Non-ESSP Schools (6 trajectories)	O_1	O_2	O_3
Cohort 8: 2007-2008 3rd Graders	**Spring 2008**	**Spring 2009**	**Spring 2010**
Original ESSP Schools (3 trajectories)	$\boldsymbol{X\ O_1}$	$\boldsymbol{X\ O_2}$	$\boldsymbol{X\ O_3}$
New ESSP School (1 trajectory)	-	$\boldsymbol{X\ O_2}$	$\boldsymbol{X\ O_3}$
Non-ESSP Schools (6 trajectories)	O_1	O_2	O_3
Cohort 9: 2008-2009 3rd Graders	Spring 2009	Spring 2010	Spring 2011
ESSP Schools (4 trajectories)	O_1	O_2	O_3
Non-ESSP Schools (6 trajectories)	O_1	O_2	O_3
Cohort 10: 2009-20010 3rd Graders	Spring 2010	Spring 2011	Spring 2012
ESSP Schools (4 trajectories)	O_1	O_2	-
Non-ESSP Schools (6 trajectories)	O_1	O_2	-
Cohort 11: 2010-20011 3rd Graders	Spring 2011	Spring 2012	Spring 2013
ESSP Schools (4 trajectories)	O_1	-	-
Non-ESSP Schools (6 trajectories)	O_1	-	-

Note: Os represent end-of-grade reading proficiency rates obtained for cross-classified cohorts and schools at the end of each school year. Xs represent the ESSP MAP. Shading and bold font indicate the four treatment trajectories.

Participants

The study took place in a community of fewer than 75,000 residents in North Carolina. The school district had an enrollment of approximately 11,000 students in 2007-2008, the first year of the study. During that year, the study sample was 51% male, 58% White, 20% Black, 13% Asian American, 10% and Hispanic-Latino American, and 5% multiracial. Overall, 18% of the district's students took part in the free- or reduced-price lunch program.

Procedures

The ESSP MAP prescribes the creation of a team of school staff members, two ESSP assessments each year, a sequence of meetings and tasks related to assessment and decision making, and the implementation of selected strategies to address threats to school success. The principal of each school determined the composition of ESSP team at his or her school. The teams were responsible for administering the ESSP assessment, using school-level ESSP data to identify one to four areas of concern, reviewing best practices using the online database, and orchestrating the implementation of strategies. The teams at all schools included teachers in the grade level targeted (e.g., third-grade teachers in Year 1, fourth-grade teachers in Year 2). Additional team members varied across the schools and included principals, assistant principals, school social workers, counselors, and parents. Five team meetings were prescribed: one for planning the pretest ESSP administration; one for reviewing data and choosing areas to target; one for selecting strategies and planning their implementation; one for

planning the end-of-year ESSP administration; and one for reviewing the posttest data. Templates were available to guide each meeting and to ensure meeting goals were achieved.

Each school was given discretionary funds for each year of the study to pay for selected intervention materials, training, and personnel. During the course of the study, each of the three original schools was given an average of $15,333 per year to pay for selected interventions. The fourth school received an average of $12,000 per year. (The average was higher for the original schools because Year 1 funds were divided three ways instead of four.) Any unspent monies were rolled over to each school's ESSP MAP budget for the following school year. We did not examine the effects of the different patterns of spending over the 3 years. Research team members attended ESSP MAP meetings throughout the study. In Year 1 of the study, the research team helped organize the collection of ESSP data and led the discussion and interpretation of assessment data. The role of the research team in these functions declined each year of the study period. In Years 2 and 3, researchers continued to attend meetings even though the school teams required little to no assistance. The research team confirmed that goals of all meetings were achieved, even when teams sometimes achieved goals outside of the formal, prescribed meetings (e.g., planning the second data collection of the year occurred outside of a formal meeting).

The Appendix summarizes the social environmental concerns identified at the four intervention schools each year and the corresponding intervention strategies selected to address concerns. Because of the ESSP MAP's emphasis on a site-level decision-making process, the concerns identified and strategies selected varied across schools. Therefore, the current study examined the effects of the ESSP MAP *process* to identify and address threats to achievement at the universal or group level.

Measures

Dependent variable. The dependent variable was the trajectory of third-, fourth-, and fifth-grade reading proficiency rates for 102 school and cohort groups. Unlike nested data in which each lower level unit is nested within an upper level unit, our clusters were cross-classified. That is, cohorts could be represented in multiple schools, and schools could be represented in multiple cohorts. Third grade is the latest federally mandated grade level in which state standardized testing can begin. Standardized tests in North Carolina (end-of-grade tests) are categorized as being at or above grade level (i.e., proficient) or below grade level based on a threshold score. Grade-level proficiency percentages in reading and math are computed for students in each grade level in a school and for demographic subgroups. Proficiency rates, which are publicly available online (NCDPI, 2011), are central to local, state, and federal school evaluation efforts. As such, schools are likely to be highly interested in strategies that demonstrate effects on these rates.

As an example of how trajectories were constructed, the third grade reading proficiency rate for the 2000-2001 cohort was obtained from each school's spring 2001 Grade 3 data. The Grade 3 data point was the intercept of each trajectory. This cohort's fourth grade percentage proficient was obtained from each school's spring 2002 Grade 4 data; similarly, the fifth-grade percentage proficient was obtained from each school's spring 2003 Grade 5 data. In addition to examining the overall percentage of students who were proficient in reading each year, we created five other outcome trajectories to represent the percentage of proficient Black students, Latino students, White students, free-lunch program participants, and nonfree-lunch program participants in each cross-classified group.

Independent variables.

Time. Time was a time-varying categorical variable and was coded 0 for third grade, 1 for fourth grade, and 2 for fifth grade (regardless of the calendar year a cohort was in any of these grades).

Treatment. Treatment was a time-varying dummy variable that took on a value of 0 at each time point (third, fourth, or fifth grade) for cross-classified schools and cohorts that did not receive the ESSP MAP (i.e., 98 of the trajectories). For the 2007-2008 cohort of third graders in the four ESSP MAP schools, a value of 1 was assigned for the treatment variable at each trajectory time point.

Controls. School-level poverty was a time-varying control variable reflecting the percentage of students participating in the free-lunch program during the years corresponding to each cross-classified unit's outcome data points. The free-lunch variable was grand-mean centered so that regression coefficients for other variables in models pertained to schools with average rates of free-lunch participation. Because the cutoff for reading scores that were considered proficient increased (i.e., became more stringent) during the first year of the ESSP MAP implementation, we included a dummy variable to represent unique intracohort effects that could be attributed to the recalibration (and other unknown cohort-specific factors).

Interactions. Two-way interactions between all combinations of time, free-lunch participation, and treatment were tested using product terms. Estimates

for treatment-X-time (grade level) product terms represented the slope effects of treatment (i.e., how the ESSP MAP moderated the effects of time on school-level reading proficiency rates). The moderation effects were of prime interest in the current study. A fourth interaction term controlled for possible slope effects associated with lower intercepts due to the recalibration of reading scores during the same year that implementation began. The term was the product of time and the cohort dummy variable described above.

Analyses

Longitudinal hierarchical linear modeling with crossed-classified random effects was used for hypothesis testing. The "xtmixed" command in Stata 11.0, (StataCorp, 2011) was used, with maximum likelihood estimation (mle), and code provided by Rabe-Hesketh and Skrondal (2005) to properly specify the cross-classified, "two-way error-components model[s]" (p. 253). The special code for cross-classified models created a three-level model (time, cohort, school) in which dummy variables were automatically created to represent each school at Level 3. With the dummy classification of schools, every combination of school and cohort was possible. The covariance matrix of the Level 2 and Level 3 random effects is specified as an identity matrix in the Stata cross-classification code. Similar to the effects of ignoring nested data structures, ignoring the cross-classified nature of data leads to underestimated standard errors (Meyers & Beretvas, 2006). Variance estimates for clustering units at other levels of the model may also be overestimated if the effects of upper level classification units are ignored (Meyers & Beretvas).

Analyses were conducted using a "long" file where each school and cohort combination had three rows of data. The three rows allowed time-varying predictors to have different values for each grade level. Time (third, fourth, and fifth grade), school-level poverty (school-level free-lunch participation rate during in a group's third, fourth, and fifth grade-year), the treatment variable, and the six proficiency outcomes were time-varying variables. Variables that did not change over time (i.e., school number and cohort number) had the same value in the three rows for a cohort-school group. The random effect of time was tested using the likelihood ratio test for nested models. Nonsignificant random effects were removed from models. The equation for the combined model, based on notation presented by Rabe-Hesketh and Skrondal (2005, p. 251) was:

$$Y_{ijt} = \beta_1 + \beta_2(time)_{ijt} + \beta_3(poverty)_{ijt} + \beta_4(Tx)_{ijt} + \beta_5(recalibration)_{ij} + \beta_{6\text{-}9}(interactions)_{ijt} + \zeta_{1i} + \zeta_{2j} + \epsilon_{ijt}$$

Y_{ijt} represents the percentage proficient reading in a particular school (i) and cohort (j) at any grade level (t) between third and fifth grade. Conditional on the dependent variable in a particular analysis, Y was the school-level percentage of proficient: (a) students overall, (b) Black students, (c) Latino students, (d) White students, (e), free-lunch program participants, and (f) nonfree-lunch program participants. β_1 is the intercept, or average proficiency level across all time points, schools, and cohorts. β_2 through β_4 are coefficients for the effects of the time-varying values of time, free-lunch participation, and treatment, respectively. β_5 is the coefficient for the effect of recalibrating of the reading cutoff. B_6 through β_9 are coefficients for the four interaction terms that were tested individually and retained only if significant ($\alpha < .05$). The term ζ_{1j} represents the random intercept for school (across all cohorts) and ζ_{2j} represents the random intercept for cohort (across all schools). The inclusion of these two random effects is what distinguishes this model from the more common nested model (Rabe-Hesketh & Skrondal, 2005, p. 251). The term ϵ_{ijt} represents unexplained variation across time for each of the cross classified school and cohort combinations.

Because of our small sample size (10 schools, 11 cohorts) and three time points, we tested relatively simple models. The school-level free-lunch participation rate was used to represent a number of highly correlated predictors of proficiency rates, only two-way interactions were tested, and nonsignificant interaction and random effects terms were removed from models.

Elaborated Hypothesis

The ESSP MAP enabled school staff to identify and address Tier 1 threats to achievement among a cohort of students as it progressed from third to fifth grade. Our primary interest was in ESSP MAP effects on the slope of the 3-year trend in school-level proficiency rates for the targeted cohort and schools. We hypothesized that, in spite of a history of lower third-grade reading scores in ESSP MAP schools, treatment trajectories would evidence steeper gains over time relative to the trajectories of comparison groups. Given the significantly lower average proficiency levels in ESSP schools relative to non-ESSP schools prior to the intervention (except among Latinos, see Table 3), we made no hypothesis about when or if the *level* of trajectories in the treatment condition would catch up with or surpass the level in nontreatment (comparison) groups.

Table 3

Mean Reading Proficiency Percentages and Standard Deviations for Treatment and Nontreatment Schools Over 7 Years (2000 to 2007) Before the First Intervention Year (2007-2008 School Year)

	ESSP MAP Schools Mean (*SD*)	Other Schools in District Mean (*SD*)	Difference
All Students	86.7 (4.0)	91.4 (3.3)	4.7***
Black	64.6 (15.3)	73.3 (13.4)	8.7*
Latino	66.9 (17.9)	73.1 (15.0)	6.2
White	95.2 (1.9)	96.0 (1.1)	0.8*
Free-Lunch Program Participants	64.7 (12.8)	72.0 (10.4)	7.3**
Nonfree-Lunch Program Participants	93.2 (2.8)	95.1 (1.5)	1.9***

Note. One ESSP MAP school did not exist at the time intervention was initiated.

* $p < .05$; ** $p < .01$; *** $p < .001$ according to independent sample t-tests.

Results

Unconditional Models

Table 4 presents estimates from the unconditional mean and growth models. As shown in the upper half of the table, although the values were statistically significant, little of the total variance or variance was attributable to school or cohort in the proficiency rates of White students and nonfree-lunch program participants. The high intercepts (95 out of 100) for the 3-year trends in these outcomes explain the lack of variance. Intraclass correlations describing the amount of total variance in outcomes explained by school and cohort effects were between .04 and .10 on these two outcomes.

There was more variance to explain in the aggregate proficiency rates of the other groups examined. Lower average third-grade proficiency rates made more variance possible across schools and cohorts. School and cohort explained similar amounts of variance in the average third grade proficiency rate for all students. However, cohort explained more variance in average starting proficiency rates for Blacks, Latinos, and free-lunch program participants than school. Intraclass correlations describing the effects of school on third-grade proficiency rates ranged from .08 for the rates of Black students to .22 for the rates of all students. Intraclass correlations for cohorts ranged from .16 for proficiency rates for Latinos to .24 for rates for all students.

As shown in the lower half of Column 5 on Table 4, annual increases in the percentage of proficient students ranged from 0.1 point for Whites to 4.7 points for Blacks. In these unconditional models, the mean change in proficiency percentages over time was not significant for Whites, nonfree-lunch program participants, or. Latinos. Likelihood ratio tests comparing results from the unconditional mean and growth model for each outcome indicated no significant variation existed around the slope (i.e., no random effects of time) for any of the trends in percentage proficient.

Final Models

Table 5 presents results from the final conditional model obtained for each dependent variable. For 4 of the 6 groups, the rate of participation in the free-lunch program had a strong and negative main effect on the percentage of students proficient in reading, suggesting that the free-lunch variable helped control for known differences between the intervention and comparison schools. The relatively small coefficients for the effects of poverty on proficiency rates among White and nonfree-lunch program participants suggested the performance of students in those groups was largely unaffected by variations in school-level free-lunch participation rates—at least in schools with up to the highest program participation rates in the school system (33%). The treatment-X-poverty and poverty-X-time interactions were not significant in any of the models; meaning the effects of the ESSP MAP did not differ by levels of school poverty, and poverty did not affect the slope of the trajectories, respectively.

The *reading-score recalibration* variable had a significant effect on the initial percent proficient in 5 of the 6 demographic subgroups examined (i.e., the intercept of trajectories). Specifically, for all but the White group, the introduction of a more stringent cut-off for proficiency resulted in a decrease in the initial percentage proficient. The finding suggests that White students tended to have high enough end-of-grade scores that they were unaffected by raising the proficiency threshold. Students in other subgroups whose proficiency was more marginal shifted into the nonproficient group when the threshold was raised. The recalibration also affected the rate of change in proficiency percentages for 5 of the 6 groups. The recalibration was associated with steeper slopes, or gains over time in the percentage of students who were proficient for all groups except Blacks. The general significance of the recalibration variable indicates controlling for the effects of the recalibration was important to avoid confounding the influence of recalibration with treatment effects.

Table 4

Estimates From Unconditional Third Grade to Fifth Grade Mean and Growth Models of Reading Proficiency

Outcome (% Proficient)	Var. of school intercept	Var. of cohort intercept	Residual	Intercept	B for grade (time)	*p* of B
Unconditional Means Models						
All students	.001	.001	.003	.88	--	--
Black	.003	.007	.025	.68	--	--
Latino	.004	.007	.030	.68	--	--
White	.000	.000	.001	.95	--	--
Free-Lunch Program	.004	.007	.020	.66	--	--
Nonfree-Lunch Program	.000	.000	.001	.94	--	--
Unconditional Growth Model (trend from Grade 3 to Grade 5)						
All students	.001	.001	.003	.870	.011*	.012
Black	.003	.007	.024	.642	.047***	.000
Latino	.004	.006	.030	.660	.023	.119
White	.000	.000	.001	.946	.001	.715
Free-Lunch Program	.004	.006	.019	.632	.029	.007
Nonfree-Lunch Program	.000	.000	.001	.935	.003	.129

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 5

Final Models of Effects of ESSP MAP on Cohort Trajectories of Reading Proficiency

	All Students	Black	Latino	White	Nonfree-Lunch Program	Free-Lunch Program
Intercept	.873	.654	.679	.950	.938	.646
Time (Grades 3 to 5)	.010*	.047***	.015	-.002	.001	.026*
School-level free lunch	-.459***	-.488**	-.452	-.059**	-.053	-.403*
Recalibration cohort	-.083**	-.177*	-.212*	-.019	-.034**	-.217**
Recalibration x Time	.048***	^	.112*	.020**	.030***	.075*
ESSP MAP	-.022	-.076	-.006	**-.069*** **	-.013	.007
ESSP MAP x Time	^	**.120***	^	**.034****	^	^

Note: ^ indicates the variable was not included in the final model because it was nonsignificant. Bold font indicates effects of the ESSP MAP.

* $p < .05$; ** $p < .01$; *** $p < .001$

The ESSP MAP treatment had significant effects on the trajectories of proficiency rates for two of the demographic groups examined: Black students and White students. The effects were over and above the effects of school-level free-lunch participation rates during each cohort's third through fifth-grade year and the recalibration of proficiency scores that occurred during the first year of the study.

Treatment effects for Blacks. Table 5 indicates that the percentage of Black students proficient in reading increased by an average of 4.7 points per year. For Black students in the ESSP MAP treatment group, the percentage proficient in reading increased another 12 points per year, or a total of 16.7 points per year. Figure 1 illustrates this statistically significant slope effect associated with the ESSP MAP for Black students in the treatment condition.

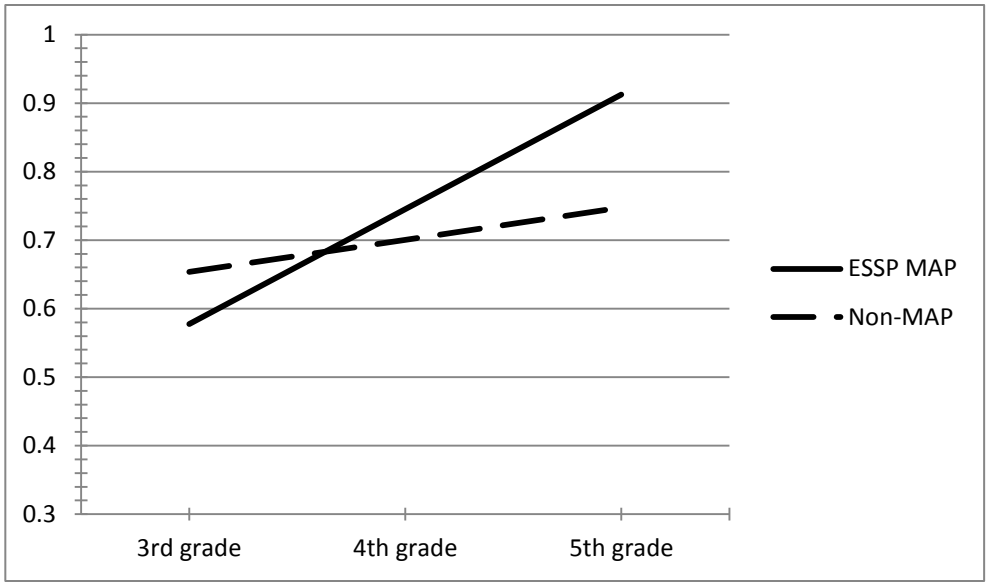

Figure 1: Predicted percentage of ESSP MAP and non-MAP students proficient in reading at each grade level: Black students. Annual gains in the percentage proficient are significantly greater in ESSP MAP schools. Trajectory levels are not significantly different at any time point.

Table 5 indicates that, after controlling for free-lunch participation rates and the recalibration, the main effects of treatment were not significant, meaning the intercepts of trajectories for Black students in the ESSP MAP condition were not statistically lower (as might have been expected) than those for their counterparts in non-ESSP MAP conditions. Because the percentage of proficient Blacks grew faster in the ESSP MAP condition than the control condition, we conducted additional analyses to determine if the main effect of the ESSP MAP was significant at the fourth or fifth-grade time points. Specifically, we changed the coding of the time variable so the fourth grade and then the fifth grade time point were equal to 0 (i.e., they were modeled as the intercept). Based on these additional analyses (results not shown in table), the coefficient for the main effect of treatment became positive and large (.12) by fifth grade (after being -.076 points at the end of third grade). However, the main effect of the ESSP MAP remained statistically nonsignificant at each time point, meaning we cannot claim the performance of Black students receiving the ESSP MAP statistically surpassed the level of proficiency of their counterparts in comparison conditions by the end of fifth grade.

Treatment effects for Whites. As shown in Table 5, the average growth in proficiency percentages among White students, controlling for other predictors was small, negative (-0.2 percentage points) and nonsignificant. However, the ESSP MAP condition affected the rate of change in this group. Among White students in the ESSP MAP condition, the percentage proficient grew by more than 3.4 points per year. As expected, the ESSP MAP was associated with a significantly lower third-grade proficiency level among Whites because low-performing schools were chosen to receive the intervention. Specifically, the average third-grade proficiency rate for Whites in the ESSP MAP condition was 6.9 points lower than that of White students in the comparison group. By repeating the analysis with fourth and then fifth grade coded as time "0," that is, the intercept, we determined that the deficit became smaller by fourth grade and smaller still and nonsignificant by fifth grade (not shown in table). In other words, as White students progressed from third to fifth grade, those in the ESSP MAP condition caught up with their non-ESSP MAP counterparts—that is, by the fifth grade, the scores of White students in the ESSP MAP condition were no longer significantly lower.

Figure 2 illustrates the slope effect of the ESSP MAP on the aggregate proficiency scores of White students. The graph indicates that White student proficiency changed little over time in the non-ESSP MAP schools and cohorts. Over and above the effects of growth associated with the recalibration (recalibration X time), the ESSP MAP was associated with annual gains in the percentage proficient. By fifth grade, the performance of White students in the ESSP MAP schools had caught up with the performance of White students in the comparison group.

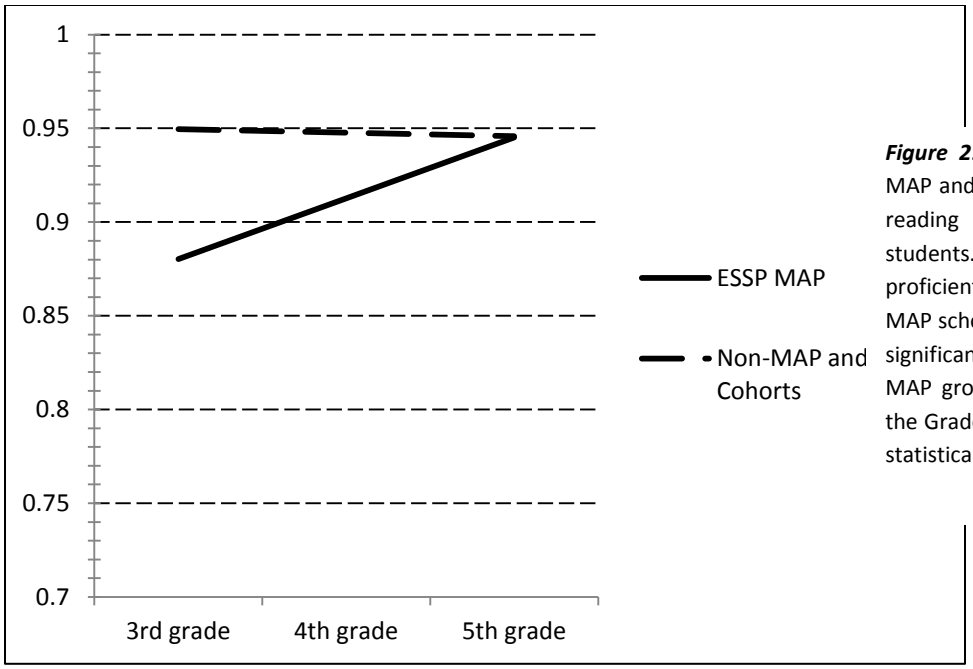

Figure 2: Predicted percentage of ESSP MAP and non-MAP students proficient in reading at each grade level: White students. Annual gains in the percentage proficient are significantly greater in ESSP MAP schools. Proficiency percentages are significantly lower for Whites in the ESSP MAP group than the non-ESSP group at the Grade 3 and Grade 4 data points, but statistically the same by Grade 5.

Treatment effects for other subgroups. Given the low performance history that led to the choice of the four ESSP MAP schools (refer again to Table 3), we expected proficiency percentages at the end of third grade (intercepts) in the ESSP MAP condition to remain lower than those in the comparison group. However, as shown in Table 5, the ESSP MAP was not associated with significantly lower starting scores for all students, Latino students, students participating in the free-lunch program, and nonfree-lunch program participants. Nevertheless, no significant slope effects were shown for the ESSP MAP for these groups. In other words, unlike the effect found for Black and White students, exposure to the ESSP MAP was not associated with more rapid increases in the percentage of proficient students for these four groups. Table 5 also indicates that the recalibration control variable had a significant effect on the slope of the aggregate trajectories for all, Latino, free-lunch program, and nonfree-lunch program students. The reduction in proficiency rates caused by the recalibration of reading scores might have made the more rapid increases in rates of proficiency possible.

Discussion

Using a quasi-experimental design, we tested the effects of the ESSP MAP on reading proficiency trajectories of 102 groups of students. The ESSP MAP provided school staff with social environmental data and supports for selecting and implementing Tier 1 (universal) interventions. Using hierarchical linear modeling with cross-classified random effects, we hypothesized the ESSP MAP would be associated with higher rates of growth in reading proficiency. To maximize the number of comparison trajectories and strengthen internal validity of the study, the comparison group included students in the same academic-year cohort but different schools, students in the same schools but different cohorts, and students in different cohorts and different schools.

Consistent with our hypothesis, the ESSP MAP was associated with greater annual gains in proficiency among two groups of students with distinctly different performance profiles in the district: Blacks and Whites. The percentage of proficient Black students in the ESSP MAP condition grew by 12 points per year, which was greater than the rate of growth for Black students in the comparison condition. The annual growth in proficiency percentages for Black students in nontreatment schools (4.7 points per year) was higher than the growth of all students and each of the subgroups examined, suggesting the district may have been working to reduce the 20-point achievement gap between White and Black students in the district as a whole prior to this study. However, the tripling of the rate of improvement evidenced in treatment schools suggests the ESSP MAP likely enhanced existing efforts. It is notable that gains of this extent were observed in schools with historically low rates of proficiency among Black students. In addition, by being associated with growth rates that far exceeded the rates achieved among White students and nonfree-lunch programs participants (in both treatment and

comparison schools), the ESSP MAP may be a strategy for reducing the achievement gap that exists between Black and White students as well as the gap that exists between low- and high-income students.

We can speculate on why the largest effects of the ESSP MAP occurred for Black students. First, ESSP data may have highlighted intervention targets that were most relevant to Black students because this subgroup was the largest nonmajority group in the school system during the study. Risk factors for Black students may have dominated the ESSP data, and therefore, the intervention goals and strategies selected by school staff. Second, it is possible that unmeasured positive changes in school staff's attitudes and behaviors toward Black students occurred as a result of learning more about their at-risk students through the ESSP data (Bowen & Powers, 2005*)*. Third, Black students in the target schools had lower starting proficiency rates than Whites and nonfree-lunch program participants, and therefore, more room to improve. Although Latinos also had low starting proficiency rates, Latinos had more variation in their proficiency rates across schools than Black students. There were also fewer data points available for Latino trajectories because of low numbers of Latinos in the school district during the early years of the study. Levels of variation may help explain the significant findings for White students: in addition to high average starting proficiency rates and little room for improvement, there was little variation in White students' rates of proficiency across schools, making it easier to detect significant differences in rates.

The percentage of proficient White students in the ESSP MAP condition increased by more than 3 points per year (3.4), a rate greater than 3 times the rate observed among their counterparts in the comparison condition. This finding suggests that White students in schools with reading proficiency rates less than 80% can benefit academically from the ESSP MAP. Unlike the percentage of Black students who were proficient at the start of the study, the percentage of White students who were proficient in reading was high (more than 90%). Differences between average proficiency rates for White students at ESSP MAP and non-ESSP MAP schools were small but statistically significant, partly due to small variances in the rates across all the schools. However, these accounts do not explain why no significant effects of the ESSP MAP were found for free- and reduced-cost lunch program participants, which is a larger at-risk subsample than Blacks and one with a smaller standard deviation in proficiency rates.

For a number of reasons, the current study might represent a conservative test of the ESSP MAP, which increases confidence in the findings. First, diffusion effects may have inflated the rates of growth in proficiency of other cohorts of students in the ESSP MAP schools during the years of the study and after. These cohorts were part of the comparison group. We did not prevent teams from including other cohorts at their schools in prevention efforts, or from choosing strategies that would have a lasting impact on classroom or school resources. The Appendix reveals many strategies that had either concurrent or subsequent effects on students beyond the cohort targeted by the study. For example, one school seeking to improve parent and school communication held a community resource information fair open to all parents. Another school purchased technology (e.g., Smartboards and audio enhancement technology) to make lessons more engaging for students with inattentive behaviors; the technology remained in classrooms after the cohort moved on, providing benefit for subsequent cohorts. Many other examples can be found in the Appendix. In addition, because all third-, then fourth-, then fifth-grade teachers at each intervention school took part in the ESSP MAP process, we expected diffusion effects on later cohorts through changes in teachers' attitudes, behaviors, and skills that may have occurred based on exposure to the ESSP MAP. Because the comparison group for the current study included earlier and later cohorts at the targeted schools who may have also benefited from the ESSP MAP, growth rates in outcomes for the comparison group may have been inflated and harder for the intervention condition to exceed.

Although the current study did not have an experimental design, the use of 11 years of data on the targeted and nontargeted schools mitigates many threats to internal validity, including history, regression to the mean, maturation, testing, instrumentation, and some selection threats. Statistically controlling for rates of free-lunch participation, which are highly correlated with school-level academic performance (Fraser et al., 2004), also helped address the known selection bias of initial performance differences between ESSP MAP schools and other schools in the district. Examination of the changes in levels of poverty in ESSP MAP and comparison schools indicated that rates were either stable (2 schools) or declined slightly in the ESSP MAP schools (2 schools, 3 points) from Year 1 to Year 3 of the study. However, rates also dropped in 4 of 6 comparison schools (declines of 1 to 5 points). Rates of poverty increased by 5 points in two of the comparison schools. With no clear pattern of change in the participation rates for the free-lunch program across conditions over the course of the study, it does not seem our positive findings for Black and White

students can be explained by free-lunch program participation rates.

The use of outcome trajectories allowed us to examine rates of change in outcomes instead of levels of performance at only one time point, which would be most strongly related to known selection bias of intervention schools. This analytic approach was even more important given that standardized testing starts at the end of third grade in the targeted state (North Carolina); therefore, we could not control for the second-grade performance of each cross-classified group. In spite of its many design strengths, a limitation of the study remains its reliance on quasi-experimental design instead of a design with random assignment. In addition, the findings likely do not generalize beyond schools in districts similar to the one studied, that is, districts in which no schools have more than 40% participation rates in free- and reduced-cost lunch programs, and in which White and nonfree-lunch program participants perform at high levels. We believe the study's reliance on a partial "effectiveness" approach to understanding the impact of the ESSP MAP on reading proficiency rates is both a strength and a weakness: the study was conducted amid real-world limitations that school personnel face on a daily basis.

Although the study was not implemented using a tightly controlled efficacy approach, it was also not fully an effectiveness study: the ESSP MAP was not implemented exactly as it would have been by a school system using the assessment and prevention model independently. Researchers initiated the effort and were present to guide and monitor the process over the 3 years of the study. In addition, our provision of incentives to parent and teacher respondents undoubtedly facilitated the collection of ESSP data, and our provision of ESSP team member incentives might have improved the investment of school staff in the process. In addition, the availability of discretionary funding for ESSP MAP teams clearly aided the process. School teams rarely, if ever, have access to discretionary monies to purchase supplies, equipment, and activities that support Tier 1 services to students. However, the amounts provided school teams in this study were not large in relation to the budgets district student services budgets. Decisions about funding school teams to carry out the ESSP MAP process would be beyond the control of most school-level personnel. School boards and district administrators would have to see the value of improving current tiered-response efforts by introducing social-environmental assessments and school-level decision making and spending authority with regard to strategies tailored to student needs.

The variability across schools and years in the combinations of interventions implemented represents an evaluation difficulty that has been encountered in other programs that appropriately (we believe) remain flexible and assessment-driven (e.g., Fast Track: Collins, Murphy, & Bierman, 2004; Multisystemic Therapy: Henggeler & Borduin, 1995). As an approach for schools to use as they respond to the idiosyncratic and temporal needs of their student bodies, the ESSP MAP is a process or a framework rather than an end product or a specific intervention. This feature is both a practice strength and an evaluation limitation.

As described above, the ESSP MAP represents a different approach to improving overall student academic performance and reducing performance gaps among students with different background characteristics than the approach that has evolved in response to NCLB (2001) and IDEA (2004) legislation. Unlike the legislative mandates, the ESSP MAP provides a social environmental assessment, access to information on best practices, and a variety of resources and supports to build skills of the school staff and make appropriate interventions feasible. The feasibility of the ESSP MAP was clearly demonstrated in this study, with the school staff at each study site successfully planning and carrying out the ESSP assessment six times over the course of the study. Each year, the school staff interpreted the data assessing social, behavioral, and environmental risk factors; prioritized concerns; and chose empirically supported interventions to buffer those concerns.

Further studies are needed to replicate the preliminary promising findings of this study. Future studies should examine the effects of the ESSP MAP (or similar approaches) in districts similar to the target district as well as districts with lower overall performance levels than the above-average system that was targeted in the current study. Researchers should also examine the process and effects in schools and districts with different student populations, including populations with a higher percentage of Latinos and Native Americans, higher proportions of students of color in general, and with more students from lower income families. Examination of the effects of the ESSP MAP on exceptional children's referral patterns and other outcomes of interest to schools is also warranted.

Although tiered-prevention frameworks that have emerged in the wake of NCLB (2001) and IDEA (2004) include best practice notions on paper, in practice, those frameworks do not provide schools with the type of data and supports necessary to properly implement tiered-response models at the local level (Alonzo

et al., 2008; Christ & Hintze, 2007; Kovaleski, 2007). Most current approaches focus on assessing the nature and magnitude of current performance problems, resulting in the overuse of secondary or tertiary efforts directly related to the instruction and learning of individual students. Such approaches ignore the less intensive, most cost-effective level of proposed tiered responses to academic and social problems. As a result, the academic achievement of students remains unacceptably low a decade after NCLB became law (NCES, 2011b). This study offers preliminary evidence that the supports and resources provided by the ESSP MAP approach partly address major shortcomings of current efforts to increase the academic performance of students in American schools though a tiered prevention framework.

Authors Note

Natasha K. Bowen is an associate professor in the School of Social Work, University of North Carolina at Chapel Hill.

Aaron M. Thompson is an assistant professor in the School of Social Work, University of Missouri: thompsonaa@missouri.edu

Joelle D. Powers is the MSW Coordinator at the School of Social Work, Boise State University: joellepowers@boisestate.edu

Correspondence regarding this article should be sent to Dr. Natasha Bowen, 325 Pittsboro St. CB #3550 Chapel Hill, NC 27599-3550, or via e-mail at nbowen@email.unc.edu

Acknowledgments

This study was supported by funding from Strowd Roses Foundation, Chapel Hill, NC and the Triangle Community Foundation. The ESSP MAP has been developed in collaboration with Flying Bridge Technologies with grants from the National Institute on Drug Abuse (1 R42 DA13865-012, 3 R41 DA13865-01S1, 2 R42DA013865-02, 1 R43DA020217-01, and 2 R44DA020217-02); the University of North Carolina at Chapel Hill (UNC); and Jane A. Pfouts and Armfield-Reeves Innovation grants from the School of Social Work, UNC. The authors have sole responsibility for the content of this manuscript.

References

Alonzo, J., Tindal, G., & Robinson, Q. (2008). Using school-wide response to intervention to close the achievement gap in reading. *Educational Research Service Spectrum*, *26*, 1-9.

Batsche, G., Elliott, J., Graden, J. L., Grimes, J., Kovaleski, J. G., Prasse, D. …Tilly, W. D. (2005). *Response to intervention: Policy considerations and implementation.* Alexandria, VA: National Association of State Directors of Special Education. Retrieved from http://www.centeroninstruction.org/response-to-intervention-policy-considerations-and-implementation

Bowen, N. K. (2011). Child-report data and assessment of the social environment in schools. *Research on Social Work Practice, 21,* pp. 476 - 486. doi:10.1177/1049731510391675

Bowen, N. K., Bowen, G. L., & Woolley, M. E. (2004). Constructing and validating assessment tools for school-based practitioners: The Elementary School Success Profile. In A. R. Roberts & K. Yeager (Eds.), *Evidence-based practice manual: Research and outcome measures in health and human services* (pp. 509-517). New York: Oxford University Press.

Bowen, N. K., & Powers, J. D. (2005). Knowledge gaps among school staff and the role of high quality ecological assessments in schools. *Research on Social Work Practice*, *15*, 491-500.

Bowen, N. K., & Powers, J. D. (2011). The Elementary School Success Profile Model of Assessment and Prevention: Balancing effective practice standards and feasibility. *School Social Work Journal, 35*, 1-15.

Bowen, G. L., Richman, J. M., & Bowen, N. K. (2002). The School Success Profile: A results management approach to assessment and intervention planning. In A. R. Roberts & G. J. Greene (Eds.), *Social workers' desk reference* (pp. 787-793). New York, NY: Oxford University Press.

Bowen, G. L., Rose, R. A., & Bowen, N. K. (2005). *The reliability and validity of the School Success Profile*. Philadelphia, PA: Xlibris Press.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.

Bryk, A. S., Raudenbush, S. W., & Ponisciak, S. (2004). *A value-added model for assessing improvements in school productivity: Results from the Washington, DC Public Schools and an analysis of their statistical conclusion validity.* Chicago, IL: Consortium on Chicago School Research.

Burns, M. K., & Gibbons, K. A. (2008). *Implementing response-to-intervention in elementary and secondary schools*. New York, NY: Routledge.

Case, A., & Katz, L. (1991). *The company you keep: The effects of family and neighborhood on disadvantaged youths.* (Working Paper No. 3705). Cambridge, MA: National Bureau of Economic Research. Retrieved from

http://www.nber.org/papers/w3705.pdf?new_window=1

Christ, T. J., & Hintze, J. M. (2007). Psychometric considerations when evaluating response to intervention. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.). *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 93-105). New York, NY: Springer.

Collins, L. M., Murphy, S. A., & Bierman, K. L. (2004). A conceptual framework for adaptive preventive interventions. *Prevention Science, 5*, 185-196. doi:10.1023/B:PREV.0000037641.26017.00

Fraser, M. W., Kirby, L. D., & Smokowski, P. R. (2004). Risk and resilience in childhood. In M. W. Fraser (Ed.), *Risk and resilience in childhood: An ecological perspective* (2nd ed., pp. 13-66.). Washington, DC: NASW Press.

Frey, A. J., Lingo, A., & Nelson, C. M. (2008). Positive behavior support: A call for leadership. *Children & Schools*, *30*, 5-14. doi:10.1093/cs/30.1.5

Fuchs, D., & Fuchs, L. S. (2005). Responsiveness-to-intervention: A blueprint for practitioners, policymakers, and parents. *Teaching Exceptional Children, 38*, 57-61. Retrieved from http://www.advocacyinstitute.org/resources/TEC_RtIblueprint.pdf

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for learning disabilities construct. *Learning Disabilities Research and Practice, 18,* 157-172. doi:10.1111/1540-5826.00072

Gordon, R. (1983). An operational classification of disease prevention. *Public Health Reports, 98*, 107-109. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1424415/

Gresham, F. M. (2005). Response to intervention: An alternative means for identifying students as emotionally disturbed. *Education and Treatment of Children, 28*, 328-244.

Henggeler, S. W., & Borduin, C. M.. (1990). *Family therapy and beyond*. Pacific Grove, CA, Brooks/Cole.

Henggeler, S. W., & Borduin, C. M. (1995). Multisystemic treatment of serious juvenile offenders and their families. In I. M. Schwartz & P. AuClaire (Eds.), *Home-based services for troubled children* (pp. 113-130). Lincoln: University of Nebraska Press.

Individuals With Disabilities Education Act. 20 U.S.C. § 1400. (2004).

Kovaleski, J. F. (2007). Potential pitfalls of response to intervention. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *The handbook of response to intervention; The science and practice of assessment and intervention* (pp.80-89). New York, NY: Springer.

Kratochwill, T. R., Clements, M. A., & Kalymon, K. M. (2007). Response to intervention: Conceptual and methodological issues in implementation. In S.R. Jimerson, M. K. Burns, & A. M. VanDerHeyden, A. M. (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 25-52). New York, NY: Springer.

Lochman, J. E., Coie, J. D., Underwood, M. K., & Terry, R. (1993). Effectiveness of a social relations intervention program for aggressive and nonaggressive, rejected children. *Journal of Consulting and Clinical Psychology, 61,* 1053-1058. doi:10.1037//0022-006X.61.6.1053

Luo, M. (2008). Structural equation modeling for high school principals' data-driven decision making: An analysis of information use environments. *Education Administration Quarterly, 44*, 603-634. doi:10.1177/0013161X08321506

Malecki, C. K., & Demaray, M. K. (2007). Social behavior assessment and response to intervention. In S. Jimerson, A. M. VanDerHeyden & M. K. Murns (Eds.), *Handbook of response to intervention* (pp. 161-171). New York, NY: Springer Science/Business Media. doi:10.1007/978-0-387-49053-3_12

Malecki, C. K., & Elliot, J. N. (2002). Children's social behaviors as predicators of academic achievement: A longitudinal analysis. *School Psychology Quarterly, 17,* 1-23. doi:10.1521/scpq.17.1.1.19902

Meyers, J. L., & Beretvas, N. S. (2006). The impact of inappropriate modeling of cross-classified data structures. *Multivariate Behavioral Research*, *41*, 473-497. doi:10.1207/s15327906mbr4104_3

National Center for Education Statistics. (2011a). *The NAEP glossary of terms.* Retrieved from http://nces.ed.gov/nationsreportcard/glossary.asp#proficient

National Center for Education Statistics. (2011b). *The nation's report card: Reading 2011* (NCES 2012–457). Retrieved from http://nces.ed.gov/nationsreportcard/pubs/main2011/2012457.apdf

National Center on Response to Intervention. (2010). *Screening tools chart.* Retrieved from http://www.rti4success.org/screeningTools

No Child Left Behind Act of 2001. P.L. 107-110, 115, Stat. 1425. (2001).

North Carolina Department of Public Instruction. (2011). *School report cards*. Retrieved from http://www.ncreportcards.org/src/

Olweus, D., S. Limber, & Mihalic, S. F. (1999). *Blueprints for violence prevention, Book Nine: Bullying Prevention Program*. Boulder, CO: Center for the Study and Prevention of Violence. Retrieved from http://www.colorado.edu/cspv/blueprints/

Ponisciak, S. M., & Bryk A. (2005). Value-added analysis of the Chicago public schools: An application of hierarchical models. In R. Lissitz (Ed.), *Value-added modeling: Issues with theory and applications* (pp. 40-81). Maple Grove, MN: Journal of Applied Measurement Press.

Powers, J. D., Bowen, N. K., & Bowen, G. L. (2011). Supporting evidence-based practice in schools with an online database of best practices. *Children & Schools*, *33,* 119-128.

Pritchard, M. E., & Wilson, G. S. (2003). Using emotional and social factors to predict student success. *Journal of College Student Development, 44,* 18-28. doi:10.1353/csd.2003.0008

Rabe-Hesketh, S., & Skrondal, A. (2005). *Multilevel and longitudinal modeling using Stata*. College Station, TX: Stata Press.

Richman, J. M., Bowen, G. L., & Woolley, M. E. (2004). School failure: An eco-interactional developmental perspective. In M. W. Fraser (Ed.), *Risk and resilience in childhood: An ecological perspective* (2nd ed., pp. 133-160). Washington DC: NASW Press.

Sameroff, A. J. (2000). Ecological perspectives on developmental risk. In J. D. Osofsky & H. E. Fitzgerald (Eds.), *WAIMH handbook of infant mental health: Infant mental health in groups at high risk* (Vol. 4, pp. 1-33). New York, NY: John Wiley & Sons.

Scott, T. M., & Martinek, G. (2006). Coaching positive behavior support in school settings: Tactics and data-based decision-making. *Journal of Positive Behavior Interventions, 8,* 165-173. doi:10.1177/10983007060080030501

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* New York, NY: Houghton Mifflin.

Shen, J., & Cooley, V. E. (2008). Critical issues in using data for decision-making. *International Journal of Leadership in Education, 11*, 319-329. doi:10.1080/13603120701721839

Spicker, H., Southern, T., & Davis, B. (1987). The rural gifted child. *Gifted Child Quarterly, 31*, 155-157. doi:10.1177/001698628703100404

StataCorp. (2011). *Stata, version 11.0.* College Station, TX. Retrieved from http://www.stata.com

Sugai, G., & Horner, R. H. (2009). Responsiveness-to-Intervention and school-wide positive behavior supports: Integration of multi-tiered system approaches. *Exceptionality, 17*, 223-227. doi:10.1080/09362830903235375

Webster-Stratton, C. (1997). From parent training to community building. *Families in Society 78*, 156-171. Retrieved from http://www.incredibleyears.com/library/items/from-parent-training-to-community_97.pdf

Webber, K. C., Rizo, C. F., & Bowen, N. K. (2011). Confirmatory factor analysis of the ESSP for Teachers. *Research on Social Work Practice.* doi:10.1177/1049731511415549

Wegmann, K. M., Thompson, A. M., & Bowen, N. K. (2011). A confirmatory factor analysis of home environment and home social behavior data from the ESSP for Families. *Social Work Research, 35*, 117-127. doi:10.1093/swr/35.2.117

Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive student behavior: Update of a meta-analysis. *American Journal of Preventative Medicine*, *33,* S130-S143. doi:10.1016/j.amepre.2007.04.011

Submitted: July 3, 2012
Revision submitted: August 31, 2012
Accepted: September 14, 2012
Published: October 9, 2012

APA citation for this article:
Bowen, N. K., Thompson, A. M., & Powers, J. D. (2012). A quasi-experimental test of the Elementary School Success Profile model of assessment and prevention. *Journal of the Society for Social Work and Research, 3,* 178-196. doi:10.5243/jsswr.2012.12

Appendix

Concerns Identified by ESSP MAP Team Each Year of the Study and Interventions Implemented to Address Concerns

	Year 1 (2007-2008)		Year 2 (2008-2009)		Year 3 (2009-2010)	
	Concerns	**Interventions**	**Concerns**	**Interventions**	**Concerns**	**Interventions**
School 1	Positive feelings about self	Mentorship from faculty and portfolios of success	Knows where to get support	Classroom discussions with social worker	Knows where to get support	Solution focused therapy groups
	Accepted by peers	Bullying/conflict resolution program in classrooms, walkie-talkies distributed to playground staff	Parent education involvement	Parent nights and Parenting Wisely intervention	Parent education involvement	Homework strategy newsletter
	Parent educational involvement	Staff development on building parent partnerships with diverse families	School performance	Tutoring, smart boards, audiology equipment, end-of grade prep materials	School performance	Smart boards, audiology equipment, end-of-grade prep materials, laptops
	School performance	Portfolios of success, tutoring, new books with fun educational activities provided to teachers				
School 2	School a fun place to learn	New student orientation, reading incentives	School a fun place to learn	After-school social clubs, affinity groups, lunch bunch groups	School a fun place to learn	After-school social clubs, affinity groups, lunch bunch groups, whole grade field trips, racial equity professional development for staff
	Uses good social skills	Anger management and social skills groups, PBIS coach hired	Parent Educational involvement	Parent nights focused on homework		
	Parent Educational involvement	Home visits and Parent nights	Working at grade level (math, reading)	During and after school tutoring	Parent Educational involvement	Parent nights
	Working at grade level (math, reading)	Tutoring and reading incentives			Working at grade level (math, reading)	During and after school tutoring
School 3	School a fun place to learn	Afterschool fun day to supplement tutoring	School a fun place to learn	Afterschool fun day to supplement tutoring	School a fun place to learn	Afterschool fun day to supplement tutoring, more family nights
	Accepted by peers	PBIS training for bus drivers, student incentives for students	Social behavior at home and school	School-wide PBIS, social skills interventions, playground incentives	Social behavior at home and school	Staff development on cooperative learning methods
	Parent education involvement	Information fair, parent night, family dance, homework books given to parents	School performance	Reward system for attendance, tutoring, math activity kits	School performance	Homework clubs, teacher training on effective instruction
	Working at grade level (reading, math)	After school tutoring				

School 4	N/A*	N/A	Knows where to get support Interacts peacefully Sociable with other children Working at grade level (reading, math)	Team building program and classroom-level interventions on social trust Steps to Respect bullying intervention In-school clubs Tutoring, taming test anxiety, growth mindset training for faculty	Good adjustment/Knows where to get support Uses good social skills Parent education involvement Working at grade level (reading and math)	Individual and small group counseling, classroom strategies on social trust and mattering SS Grin social skills program and leadership program Parent resource library and Parent nights at school Tutoring, class visits by college students, growth mindset for faculty

Note: Target school #4 opened in the 2008-09 academic year and was not involved in the first year of the study.