10-21-2016

# Academic and Psychosocial Outcomes of a Physical Activity Program with Fourth Graders: Variations Among Schools in Six Urban School Districts

Cheryl L. Somers
*Wayne State University*, c.somers@wayne.edu

Erin E. Centeio
*Wayne State University*, erin.centeio@wayne.edu

Noel Kulik
*Wayne State University*, ab7564@wayne.edu

Alex Garn
*Louisiana State University*

Jeffry Martin
*Wayne State University*, aa3975@wayne.edu

***See next page for additional authors***

## Recommended Citation

Somers, C. L., Centeio, E. E., Kulik, N., Garn, A., Martin, J., Shen, B.,... McCaughtry, N. A. (2016). Academic and psychosocial outcomes of a physical activity program with fourth graders: Variations among schools in six urban school districts. *Urban Education*. doi: 10.1177/0042085916668951
Available at: http://digitalcommons.wayne.edu/coe_khs/71

**Authors**
Cheryl L. Somers, Erin E. Centeio, Noel Kulik, Alex Garn, Jeffry Martin, Bo Shen, Mariane Fahlman, and Nathan A. McCaughtry

Academic and psychosocial outcomes of a physical activity program with fourth graders: Variations among schools in six urban school districts

Cheryl L. Somers[1], Erin E. Centeio[1], Noel Kulik[1], Alex Garn[2], Jeffrey Martin[1], Bo Shen[1], Mariane Fahlman[1], & Nathan A. McCaughtry[1]

Wayne State University[1]

Louisiana State University[2]

We would like to thank Blue Cross Blue Shield of Michigan, the Blue Cross Blue Shield of Michigan Foundation, and the United Dairy Industry of Michigan for their generous financial support of this research.

## Abstract

The purpose was to examine academic achievement, school attachment, and peer acceptance before and after a comprehensive school-based physical activity program (CSPAP) with 378 children in 12 fourth grade classrooms across six schools in primarily low SES districts of a large Midwestern metropolitan area. Both personal and normative rate of academic achievement improvement metrics were used. Overall, all students showed personal math and reading growth. However, effects varied by types of achievement indicator and comparison group, revealing noteworthy school-level demographic and implementation characteristics that are inextricably intertwined with program effectiveness and student growth. Implications, especially for minimizing generalizations, are significant.

## Introduction

A comprehensive school physical activity program (CSPAP) called Building Healthy Communities (BHC), which emphasized physical activity (PA) education and opportunities, was implemented during this study, with a primary goal of improving school day PA among students. An urban context was selected for this program implementation and research because obesity, lower activity levels, and poorer nutrition tend to be significantly higher in urban poor areas than in typical suburban, middle SES environments, and America's urban centers, as described as primarily within major cities and metropolitan areas, have disproportionately high poverty, financial challenges, low academic achievement, rampant education disparities, crime exposures, etc. (Hodge & Vigo-Valentin, 2014). Hodge and Vigo-Valentin (2014) also described the discrepancy between what the research community and government agencies profess as best for students regarding physical activity and nutrition, on the one hand, and the fact that it so infrequently happens fully for children in urban contexts, on the other.

This can be attributed to both structural systemic inequities that contribute to academic underperformance by the almost 7 million students living in poverty (Milner & Lomotey, 2014). Importantly, predictors of obesity have been conceptualized as situated in two categories—modifiable and nonmodifiable (Hodge & Vigo-Valentin, 2014) and as embedded in variables occurring both inside and outside of the school environment (Milner, 2013). In school, for example, we can more easily impact teaching practices and administrative decision-making than we can school funding. At home, we can disseminate to parents information about how to help their children exercise more but impacting their finances to buy healthier but costlier foods and to change major habits is more difficult. School-based interventions/programs have historically involved micro-level practices that are inadequate to meet the needs of students living and attending schools in urban contexts (Milner & Lomotey, 2014). Given this, the CSPAP utilized in this research was designed to target a wide range of modifiable predictors in the school day context such as physical activity, general movement, and eating behavior, and to extend further to impact students' not only in their physical activity and nutrition behaviors, but in their health self-perceptions, academic achievements, and several areas of social-emotional functioning.

CSPAPs give children the opportunity to meet the national recommendation for 60 minutes of daily exercise (Donnelly et al., 2008; Sallis et al., 1999). The BHC program emphasized implementation of multiple components of a CSPAP (AAHPERD, 2013), including recommendations for physical fitness interventions (e.g., Hodge & Vigo-Valentin, 2014) and early research has shown its effectiveness. For example, Centeio et al. (2014) reported that BHC significantly improved children's moderate-to-vigorous physical activity, as well as step counts, which were the most proximal outcomes expected from the program (Centeio et al., under review; Centeio et al., 2014).

PA has been encouraged throughout the school day because there is a relationship between students' PA and important academic achievement (AA) outcomes. A recent systematic review of four decades of studies (see Castelli et al., 2014) revealed fairly convincing evidence that PA is linked to both cognitive functioning (e.g., working memory, processing speed, verbal and spatial reasoning) and AA (math, reading), and PA interventions have been shown to positively impact both cognition and AA. Effect sizes reflecting magnitude of impact indicate that the effects a few decades ago were small (ES=.212) but have become stronger in the past two decades (ES= .496), perhaps due to measurement or methodological evolution, or to advances in the quality of intervention design and implementation. The overall effect size across all decades of studies was .383. Other researchers conducting comprehensive reviews have reported similar findings (Biddle & Assare, 2011; CDC, 2010; Donnelly et al., 2009; Singh, et al., 2012; Trudeau & Shepard, 2008).

These relationships between PA and AA have been explicitly shown in African American female adolescents (e.g., Shen, 2015).

Given that CSPAPs may improve PA, and that PA is indeed important for AA, the goal of this study was to examine the links between the BHC program and more distal variables like academic achievement and psychosocial factors. It was hypothesized that this CSPAP would improve the AA in the urban/metropolitan, primarily minority elementary students included in this program. Also, as indicated above, much prior research has emphasized cognition outcomes and results have varied. Cognition underlies achievement but explains no more than about half and often less of the variance in AA (e.g., Deary, Strand, Smith, & Fernandes, 2007; Duckworth & Seligman, 2005; Naglieri & Bornstein, 2003). Furthermore, academic skills are a critical precursor to remaining in school, which predicts greater occupational earnings and other successes (Steinberg, 2014). Therefore, more research explicitly targeting academic achievement through PA programs is warranted. Rigorous approaches to measuring academic achievement, specifically reading and math skill development, was the focus in the current study. Specifically, curriculum-based measures (CBMs) were used, which are more closely tied to instruction, more sensitive to change over time, and psychometrically equivalent or superior to more global AA measures like grades or standardized test scores (Deno, 2003; Fore, Burke, & Martin, 2006; Good & Jefferson, 1998; National Center on Student Progress Monitoring, 2007; Shinn, 1998; Shinn & Shinn, 2002).

Additionally, social-emotional functioning and social interaction were anticipated to be important correlates of the development of physical fitness skills and academic skills alike. Thus, school attachment and peer acceptance were included as potential correlates because attachment to school has been shown to be positively related to academic engagement and motivation as well as academic success expectancies (Goodenow & Grady, 1993). Others have found that when the social environment is enhanced, as was expected to happen in the BHC program, students bond to school more, excel in academics better, and have fewer behavior problems (Catalano et al., 2004). Other factors like social support in school, e.g., how well students feel connected to and cared for by teachers and others at school (Wang & Eccles, 2012), and the social context of the classroom, e.g., how socially accepted and connected students feel (Baker, 1999), were both associated with more satisfaction with school. All of this is likely related to being attached to school, and less school attachment is related to a pattern of negative school behaviors that can lead to failure and drop out (Voelkl, 1996). Therefore, because the secondary goal of this research was to improve achievement and keep student engaged in school, school attachment was included as a measure. As teasing and poor peer relations are related to both obesity and academic achievement (e.g., Gunnarsdottir, Njardvik, Olafsdottir, Craighead, & Bjarnason, 2012), perceived peer acceptance was included similarly for its association with the broader school climate and context, and to target whether the children perceived improvements in peer acceptance over the program year. Some researchers have also combined both variables and found that school connectedness, as well as peer harassment, was negatively associated with academic achievement (e.g., Eisenberg, Neumark-Sztainer, & Perry, 2009).

From a Social Ecological Framework (SEF) perspective, multiple factors from multiple life contexts/levels of influence work together to influence health behaviors, and both the physical and social environments have been implicated by Sallis, Owen, and Fisher (2008) and Stokols, (1992). This framework is drawn from Urie Bronfenbrenner's bio-ecological theory (1977; Bronfenbrenner & Morris, 1998, 2006), which extends from the classical dyad (parent-child) to consider the overall context of a child's growth. The individual's intrapersonal variables are at the center of one's developmental structure. The intermediate setting of a person is called the

*Microsystem* (e.g., exposures to/influences of family, immediate home environment, school, peers, etc.). Next is the *Mesosystem*, which refers to the interactions between various microsystem contexts (e.g., parent-teacher communication and relationships, parent-sibling dynamics, etc., that impact the individual child). The *Exosytem* refers to factors that do not directly impact children but that impact them indirectly through others (e.g., parents' job dynamics that parents bring home). Our overarching cultural ideologies, norms, practices, etc., involving politics, religion, economics, etc. are part of the *Macrosystem* and can have impact on individuals. There is also a *Chronosystem*, which refers to the impact of change itself over time. These systems are interconnected and produce varying degrees of impact on children as they develop. Research tapping as many of these complex layers as possible is ideal.

In light of this, as well as the urban education context discussion earlier, the associations that were expected in this current CSPAP/project are logical through this lens, as well as when considered through a biopsychosocial lens. This is especially so when considering the earlier discussion about the impact of poverty on access to healthy models and options (e.g., Hodge & Vigo-Ventalin, 2014) and the fact that obesity has been suggested to be a proxy for poverty (Katz, 2013). In all children, one's biological/physical development (e.g., physical health, obesity, etc.) is not independent of one's social-emotional and cognitive or academic development; there are interactions that cannot be ignored (e.g., Baltes, Reese, and Lipsitt, 1980). Similarly, one's attachment to school and acceptance by peers is likely impacted by a combination of physical and emotional factors in the environment, such as assumptions that teachers make, both positive and negative, about poor students' capacities to learn (Biag, 2016)/ Teachers' preparedness to educate children from impoverished backgrounds also likely contributes to the ways in which teachers and students interact and the ways that students feel about being at school (Ullucci & Howard, 2015).

Other developmental researchers (e.g., Sameroff, 2010) similarly recognize the interconnectedness of the individual and his/her life contexts to explain development, much as did scholars like Bronfenbrenner (1979) through his bioecological model of human development. To examine these layers simultaneously reflects a broader recognition of the complexities of development.

Thus, in consideration of a child's full ecology as well as the need to permeate the health behavior of children more comprehensively through school-based education and intervention efforts, the overarching goal of this project was to provide students, teachers, and administrators with professional development and resources to transform the elementary school environments for a daily focus on PA. These tools included education about PA, exposure and access to activity programs, opportunities for regular PA breaks throughout the day, and fitness/sports equipment with which to play and be active. These components were designed to permeate the daily schedule, affect everyone involved, and thereby alter the school climate and culture around health and fitness (http://coe.wayne.edu/centerforschoolhealth/programs.php). This program was implemented and studied for effectiveness to address gaps in the literature.

**Purpose of the Study**

Based on this theoretical grounding in the interaction between biological and environmental influences, as well as ecological models of human development and behavior change, the purpose of this study was to extend prior studies through these research aims: 1) to examine personal growth in academic achievement over one school year, 2) to examine participants' academic achievement outcomes compared to national norm sample growth over the school year before and after the program, and 3) to examine the peer acceptance and school attachment as important potential program outcomes. Overall, it was expected that the PA program would lead to significant improvements in academic achievement, peer acceptance, and school attachment. Disadvantaged, urban and some urban/metropolitan students were of focus in this work due to rising rates of poor exercise and eating habits/nutrition especially among African-American samples.

## Method

### Participants

The sample of 378 boys (n=209) and girls (n=169) was taken from 4th grade classrooms (14 classrooms total; average age=9) at six schools in urban areas of a large Midwestern city. The participants were primarily African American (n=172; 46%) or Caucasian (n=98; 26%), with bi-racial the next most common category (n=53, 14%). Free and Reduced Price School Meal rates were moderate to high (50-89%) in all but one school, which was 19%.

### Measures

In addition to demographic questions, the following measures were also taken from the children's perspectives. This was done at both time 1 (T1; pre-program) and time 2 (T2; post-program).

**Academic achievement.** The Academic Improvement Monitoring System (AIMSweb; www.aimsweb.com) and the DIBELS (Dynamic Indicators of Basic Early Literacy Skills) are two different sets of brief, direct measures of academic skills commonly used in K-12 schools as universal screening tools to identify those not attaining benchmark skill levels and to monitor those who are significantly below grade level and in need of remediation. They are both called curriculum-based measures (CBMs) and are universal in that the skills tested are generally considered to be consistent benchmarks across school buildings, districts, and states, representing another hallmark of this methodology. There are many equivalent versions all calibrated to the same difficulty level within each grade level. They are sensitive to change over time and in response to interventions. Raw scores in math computation (using AIMSweb) and reading comprehension (using DIBELS) were obtained before and after the program. Also, the Rate of Improvement (ROI) for AIMSweb math and DIBELS reading were obtained from the national norm charts and participants' own ROIs were computed as post-test score minus pre-test score divided by the approximately 30 weeks in the program, which corresponds to approximately the time between fall and spring benchmarks for the national sample.

The reliability and validity of CBMs has been well established (Deno, 2003; Fore, Burke & Martin, 2006; Shinn, 1998; Shinn & Shinn, 2002). Specifically, reading CBMs compared with global reading skills have coefficients of agreement from 0.63 to 0.90 with most exceeding 0.80 (Marston, 1989), and a recent meta-analysis of 29 studies demonstrated an overall agreement of 0.69 (Yeo, 2010). Correlations to performance on global state tests range from 0.5-0.9 and 0.6-0.9 range for math computation and math concepts/applications, respectively (Keller-Margulis, Shapiro & Hintze, 2008;Marston, 1989; Santoro, Lutz, Keller, Hintze & Shapiro, 2006; Thurber, Shinn & Smolkowski, 2002) while discriminant analyses reveal that CBM math probes predict performance on state math standards with 87% accuracy (Tindal, Helwig & Anderson, 2002). High reliability of CBMs has been shown for reading comprehension, 0.74-0.81 (Deno, Shin & Espin, 2000; Tolar, Vaughn, Stuebing & Barth, 2012).

**School attachment.** Defined as a student's overall connectedness to school, a ten-item scale developed by Somers and Gizzi (2001) was used to identify participants' level of school attachment. Sample items included "School is important in my life" and "School is one of my favorite places." Students responded on a five-point scale (1 = "Strongly Disagree," 5 = "Strongly Agree). Cronbach's alphas, a measure of internal consistency reliability were .88 (Somers and

Gizzi, 2001) and .84 (Weber, Somers, Day, & Baroni, 2016). In the current sample, the alphas were .80 and .79 at pre and post-test, respectively.

**Peer acceptance.** A five-item scale was developed to measure acceptance by peers: "Do most kids like you?," "Do you have friends?," "Do kids like you just the way you are?," "Do other kids want to include you with them?," and "Do other kids want to hang out with or play with you?" Response options ranged from 1=No, 2=sort of no, 3=not sure, 4=sort of yes, 5=yes. This was piloted prior to use in this study and was found to be easier for elementary school-aged children to understand than "strongly agree" to "strongly disagree" options. Responses were summed and averaged and totals reflected relative acceptance from peers. Alphas for the current sample were .86 at pre-test and .79 at post-test.

## Procedure

For this CSPAP, elementary schools in urban, primarily economically disadvantaged areas of the Midwestern United States participated in this whole-school intervention and research study, targeting students, educators, and parents for PA and nutrition programming. Four CSPAP elements were included: (1) quality Physical Education using innovative PE approaches, (2) classroom PA through which children were given frequent movement and other PA breaks, (3) opportunities for quality PA at lunch and recess, and (4) opportunity for an after-school PA club through which students processed healthy eating and PA more deeply. All of these components were designed to work together to change the school culture toward one in which PA and healthy eating were the norm. Staff were given regular trainings before the school year and ongoing professional development. They were encouraged to be active role models for students and to change messages about exercise and eating. They did this by modeling healthy eating themselves, talking openly about choosing healthy snacks, exercising, and being satisfied by that, praising students for healthy choices, and general positive emotion around feeling good when eating well and exercising. Materials and presentations/ trainings were offered to parents as well. Project coordinators were placed in schools routinely throughout the year to ensure fidelity of implementation of all aspects of the intervention as well as the research/data collection process.

Data collection occurred at the start and near the end of the school year. Physical activity was measured via accelerometers worn for three days at each time point, which yielded "average steps taken per day" as well as amount of time spent in moderate-to-vigorous physical activity (MVPA). The aforementioned surveys and academic testing was also administered to all students at both data points. All procedures were fully approved by the university Institutional Review Board for safe treatment of human intervention and research participants.

## Results

Means and standard deviations are included in Table 1. Analyses were begun by comparing students' reading and math scores over the 30 weeks of the program to the national norm sample growth over the same period of time. One-sample *t-tests* were conducted, with the test value being the national $4^{th}$ grade ROI of .40 reading comprehension points correct per week and .72 math computation points increase per week. The average ROI for this sample was .26 reading points per week (SD=.23) and .56 math points per week (SD=.48). These were both significantly less than the national norm sample ROIs, respectively t(281)=-5.50, p=<.001.00 and t(277)=-4.90, p=.000. Students in this sample evidenced an average ROI that was at a significantly slower rate than that of the national norm sample.

Next, the pre-post scores for the full sample were examined via paired samples *t-tests* and found that their academic achievement did improve pre to post (see Table 1) at statistically significant levels even though their pace of growth was slower than the national norm sample.

Then, peer acceptance and school attachment were analyzed (Table 1). Results indicated no significant changes in peer acceptance over time, but that school attachment decreased from pre to post. Means indicate high scores on both the peer acceptance and school attachment constructs at pre-test, and similarly relatively high scores at post-test. In an effort to understand why the aggregate school attachment variable decreased to a statistically significant degree from start to end of the school year, an ANOVA was conducted to test for school level differences. Analyses revealed there were differences in school attachment by school [$F(5,314) = 4.11$, $p < .001$] but not for peer acceptance [$F(5, 323)=.23$, $p=.948$]. There remained this school-level difference in school attachment at the end of the year as well [$F(5, 285)=4.41$, $p=.001$] and again not for peer acceptance [$F(5, 292)=.62$, $p=.685$]. Thus, additional analyses were run separately for each school (Table 2) for school attachment. Results now indicated that school attachment at one school (school #3) skewed the overall data for the full sample in that in only one other school showed a substantial drop over the year in school attachment.

This prompted an exploration of academic differences in raw math and reading scores separated by school (see Table 2). Indeed, there were school-level differences for reading and math as well. In all but one school, average raw math and reading scores improved from pre to post program. Math scores improved by a few points (from 58 to 61 raw score points) at school 2, but were not statistically significant. Their reading scores did improve significantly and for all other schools, substantial gains for both reading and math were observed.

Finally, ANOVAs were run with ROI for reading and math as the dependent variables by school. There was a significant difference by school for Math ROI, with school 2 significantly lower than either schools 3, 4, and 6, though all four schools were still lower than the national ROI. All four were also significantly different than schools 1 and 5, which each evidenced ROIs faster than the national growth norm [F(5, 276)=16.98, p=.000]. School 1 is higher SES (only 19% Free and Reduced Price Meals/FRPM), but school 5 is among the lowest SES in the sample (78% FRPM). For Reading ROI, schools 2, 4, and 5 were significantly lower than schools 1, 3, and 6 [F(5, 272)=20.19, p=.000]. There is no apparent SES pattern within these results. See Table 3 for both math and reading results.

Last, one-sample *t-tests* were re-run, separated by school, comparing this sample's ROIs to the national norm sample ROIs (see Table 3). The participants' Math ROI was significantly faster than the norm sample for schools 1 [t(46)=2.61, p=.012)] and school 5 [t(38)=2.51, p=.017] was significantly lower on the ROI Math score, as was school 3 [t(72)=-4.43, p=.000], school 4 [t(37)=-2.94, p=.006], school 6 [t(60)=-5.27, p=.000], and even more profoundly slower than the norm sample for school 2 [t(24)=-7.43, p=.000)]. For Reading ROI, schools 3 [t(71)=.80, p=.429] and 6 [t(37)=.14, p=.890] were not significantly different than the national average ROI, but the other four schools made gains at a significantly slower rate than the national average: school 1 [t(47)=2.61, p=.012], school 2 [t(24)=.7.43, p=.000], school 4 [t(37)=-2.94, p=.006], and school 5 [t(38)=2.51, p=.017].

## Discussion

The purpose of this investigation was to describe the academic achievement, school attachment, and peer acceptance outcomes before and after a CSPAP. Past researchers have demonstrated an important link between PA and academic achievement. In implementing this intervention the goal was to increase students' PA through a whole-school approach, and to then to also have an impact on academic achievement, and ideally residual influences on peer acceptance and students' attachment to school, which have been likewise correlated with school success. One main finding was that when comparing children's pre-test to post-test scores within each individual, the students did grow in reading and math.

Our study design also involved a comparison of this sample to national ROI on academic success measures of math and reading comprehension. Although this sample grew academically between pre and post testing, it was another key finding that the students did not demonstrate *rates* of growth on par with the national norms curve. However, the national norm group includes all demographic backgrounds, including very high achieving schools from mid to high socioeconomic backgrounds. The fact that the lower SES sample was compared to students not necessarily representative of them is a likely explanation for this, and had these data simply been analyzed in comparison to national samples, it would appear as if there was no impact, and perhaps even a deleterious effect, as this sample grew, but at a slower rate than the national sample. Upon more careful consideration of what was assumed to be a homogeneous group of "at-risk, urban/metropolitan schools," it became clear that there was actually more heterogeneity than demographic data suggest on the surface, and that generalizations would be inappropriate.

This is a noteworthy result that has complicated implications for interpretation and future research and intervention work. It is well established by scholars that we have disproportionate problems with learning in urban contexts due to educational inequities associated with poverty (e.g., Hodge & Vigo-Valentin, 2014), yet practitioners continue to compare student performance to averages across all students. It is a critical finding here that, despite the fact that overall these collective students' ROI was slower than the national samples, when examining pre-post raw score gains there is indeed a significant gain, in some less than others due to suspected important school-level differences, but most students demonstrated substantial gains. It quickly became apparent that considering the sample in aggregate also did not allow us to uniformly detect unique gains in the subpopulations within this group of students. There were school level differences that partially seem to explain the lack of findings. With a large enough sample size, these differences could of course be controlled for statistically, but understanding the school-level differences became of greater interest. The six schools were primarily lower than, with only some on par with or higher than, the national rate of improvement. Indeed, there was a school level difference for reading, math, and school attachment for some schools that was worth exploring more deeply.

In an attempt to understand what is being reflected by these school-level differences, socioeconomic status (via Free and Reduced Price Meals--FRPM, as is noted in Tables 2 and 3), school type (public and private), and school size were available data to consider for this sample. It is also important to note that some schools started out the year with very low scores. An observational analysis suggests that there was no clear pattern in what was significantly improved or not and for which schools. In the case of Math ROI, for example, the poorest ROI was in a private school with half free and reduced price meal (FRPM) percentages and a smaller school size, yet the strongest Math ROI occurred in a public school of similar small size, yet nearly 90% FRPM. And perhaps most interesting was that the only other school to demonstrate Math ROI greater than the national ROI was also a public school of small size but with only 19% FRPM status. Those two schools had very different SES (as measured via FRPM) though similarly high outcomes compared to national standards. These are inconsistencies that prevent a clean interpretation of what happened in the sample (e.g., two schools that were not significantly different from each other on Math ROI were both same small size, but they were the second most FRPM (78%) and the lowest FRPM (19%). Thus, conclusions cannot be drawn about the impact of poverty. They were both small schools, though, and that is worth looking into further. Then in reading, the two schools that excelled most were different still than the highest math performing schools in this sample. These two schools were among the highest FRPM (77% and 81%) public schools with double the school size of the smaller schools. This is indeed consistent with prior literature that showed neighborhood differences in academic achievement among multiple impoverished neighborhoods (Kim, Mazza, Zwanziger, & Henry, 2014). However, this clearly needs more exploration in future studies, as some of these results are difficult to explain.

Therefore, relying on aggregate growth indices is not likely the ideal way to measure growth in the samples that the project aimed to affect. One of the most important conclusions from this work is that consideration of *personal* growth over the year, not only comparison to state or national standards, is critical. In fact, there is a national movement toward using individual rate of growth models in place of the high stakes norm-based comparisons because they appear to be generally more fair to students who simply start behind the average (Ballou, Sanders, & Wright, 2004; Koretz, Louis, & Hamilton, 2004; Sanders, 1998). These proponents advocate for alternative metrics to holding teachers accountable for their performance, because students who perform poorly at the beginning of the school year require more educational capital to remediate. There is solid argument for rethinking how students are assessed and compared to each other, particularly among those who are impoverished, and how teachers are assessed (Carey, 2014). This is especially poignant when considering the overrepresentation of minorities who are poor in special education, which typically results from comparisons to average scores across all economic and racial groups (Griner & Stewart, 2013). Analyzing personal growth is, therefore, in line with this movement. This is perhaps one of the most important implications of these findings—that schools must begin to seriously consider students' individual rates of growth, starting with the points at which they begin, when making major assumptions about and plans for students.

Regarding the psychosocial variables, a goal of education is generally to see increases in such factors as peer acceptance and school attachment. However, it is important to note that these students started the year with relatively high perceptions of peer acceptance and attachment to their schools, and also that with one school as the exception, students did not decrease in school attachment over the year. Students being attached and feeling accepted is contrary to what is often incorrectly assumed about urban and poor contexts and such generalization about at-risk contexts perhaps should be challenged. It is an important finding here that these students were by and large *not* detached from their schools. Implications for schools are to avoid predisposed assumptions about what students in urban contexts bring to the school environment in terms of academic orientation, and to thus hold higher expectations for the attitudes and behaviors that students will arrive at school with from home. Results indicated that they *do* feel connected to their buildings and they are liked by peers overall.

There are several issues intertwined here, the first two of which have been discussed above: a) What are the implications of findings such as these for the academic growth for these children in general? b) What does it mean for how to examine these children's progress? By what standards is success/growth to be judged? And, next, c) What does this mean for the selected PA and nutrition intervention program specifically? As with many intervention programs, it is difficult to attribute an effect to a program beyond the immediate obvious impact (e.g., in this case, actual PA). However, an improvement in students' PA was demonstrated in terms of increased average steps taken per day after participation in this program, as well as a link between PA and higher ROI for math (Centeio et al., under review). A long term goal is to continue to move toward impacting academic and psychosocial factors as well, as health and fitness is inextricably woven into academic and psychosocial functioning and motivation. The goal in the current paper was to explore more deeply the ways in which academic outcomes are measured among seemingly similar but actually different subsamples of students.

In conclusion, the children in this sample did grow over the year both academically and with physical activity. There are also lessons learned in this process of carefully evaluating outcomes of the program. Limitations include a need to match schools and comparison groups carefully so that nested models and analyses can be conducted. There are also variables not explicitly measured here, e.g., preparation for schooling and personal/family history, as these were assumed based on school socioeconomic estimates. Family commitment to achievement and level of engagement in youth development in the communities could also be included in future research as they are likely to play a salient role, possibly even irrespective of economic levels, especially as, through a Social-Ecological Framework (SEF) lens as described earlier, it is critical to consider impacts on children's development at multiple contextual levels (Sallis, Owen, & Fisher, 2008; Stokols, 1992). Also, other factors can be included in future attempts to tease out these inter-school variabilities, such as school climate for PA and academics (e.g., value placed on and models demonstrated regarding PA and academic achievement by the school), parent involvement and attitudes, instructional approaches to reading and math, and characteristics of teachers' motivational strategies, instructional patterns, and interpersonal engagement skills. These may be important school-level factors to consider. There is also a need to continue to consider children's *individual* rates of growth and not only consider group data. Additionally, although study personnel were trained consistently and together and placed them in the schools to routinely monitor fidelity of program implementation and data collection, inter-school variations could occur.

Taken together, however, these results support the notion that researchers should not generalize across seemingly similar contexts based on like demographics or a general “at-risk” classification. Controlling for demographics is often used, but the approach used in this study actually helps to make the case that considering samples across schools in aggregate misses important nuances. Qualitative data analyses also provide rich information that cannot be captured in traditional analytic methods in order to tease out subsample variations such as those highlighted in this study. All of this information helps us to better understand how to design interventions more effectively in an effort to expand the impact of these types of school-based interventions. In that vein, educational and health education professionals will benefit from this information and perspectives regarding what to keep in mind when selecting and tailoring interventions to their students, and when selecting methods of measuring impact and then making interpretations of those data. Implications are significant for the children when researchers and practitioners make mistakes in intervention implementation or overextension of results interpretation. This is particularly true in the urban settings in which these data were collected, as it can set low expectations for students from disadvantaged backgrounds as well as biases about what they can accomplish.

# References

American Alliance for Health, Physical Education, Recreation and Dance (AAHPERD). (2013). Comprehensive school physical activity programs: Helping students achieve 60 minutes of physical activity each day [Position statement]. Reston, VA: Author. [This organization is presently known as SHAPE America]

Baker, J. A. (1999). Teacher-student interaction in urban at-risk classrooms: Differential behavior, relationship quality, and student satisfaction with school. *The Elementary School Journal*, *100*(1), 57-70.

Baltes, P. B., Reese, H. W., & Lipsitt, L. P. (1980). Life-span developmental psychology. *Annual Review of Psychology, 31*, 65-110.

Ballou, E., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics, 29* (1), 37-65.

Biag, M. (2016). A descriptive analysis of school connectedness: The views of school personnel. *Urban Education, 51* (1), 32-59.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*(7), 513-531.

Bronfenbrenner, U., & Morris, P. (1998). The ecology of developmental processes. In W. Damon & R. Lerner (Eds.), *Handbook of Child Psychology: Theoretical Models of Human Development, 5th Edition* (Vol. 1, pp. 993-1028). Hoboken, NJ: John Wiley and Sons.

Bronfenbrenner, U., & Morris, P. (2006). The bioecological model of human development. In R. Lerner & W. Damon (Eds.), *Handbook of Child Psychology (6th edition): Theoretical Models of Human Development* (Vol. 1, pp. 793-828). Hoboken, NJ: John Wiley & Sons.

Carey, R. L. (2014). A cultural analysis of the achievement gap discourse: Challenging the language and labels used in the work of school reform. *Urban Education, 49* (4), 440-468.

Castelli, D. M., Centeio, E. E., Hwang, J., Barcelona, J. M., Glowacki, E. M., Calvert, H. G. and Nicksic, H. M. (2014). VII. The history of physical activity and academic performance research: Informing the future. *Monographs of the Society for Research in Child Development*, *79*, 119–148. doi: 10.1111/mono.12133

Catalano, R. F., Haggerty, K. P., Oesterle, S., Fleming, C. B., & Hawkins, J. D. (2004). The importance of bonding to school for healthy development: Findings from the social development research group. *Journal of School Health, 74*(*7*), 252-261.

Centers for Disease Control and Prevention, (CDC) (2010). *The association between school-based physical activity, including physical education, and academic performance.* Atlanta, GA: U.S. Department of Health and Human Services.

Centeio, E.E., Somers, C.L., Kulik, N., Garn, A., Shen, B., Martin, J., Fahlman, M. & McCaughtry, N. (under review). The Relationship Between Academic Achievement and Healthy School Transformations in Urban Elementary Schools.

Centeio, E.E., McCaugtry, N., Gutuskey, L., Garn, A., Somers, C., Shen, B., Martin, J., & Kulik, N. (2014). Physical activity change through comprehensive school physical activity programs in urban elementary schools. *Journal of Teaching and Physical Education, 33*(4), 573-591.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35*, 13-21.

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education, 37* (3), 184-193.

Deno S. L., Shin J., & Espin C. (2000). Technical adequacy of the maze task for curriculum-based measurement of reading growth. *The Journal of Special Education; 34*, 164.

Donnelly, J. E., Greene, J. L., Gibson, C. A., Smith, B. K., Washburn, R. A., Sullivan, D. K., . . . & Williams, S. L. (2009). Physical activity across the curriculum (PAAC): A randomized controlled trial to promote physical activity and diminish overweight and obesity in elementary school children. *Preventive Medicine, 49*(4), 336–341.

Duckworth, A. L., & Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16* (12), 939-944.

Eisenberg, M. E., Neumark-Sztainer, D., & Perry, C. L. (2003). Peer harassment, school connectedness, and academic achievement. *Journal of School Health, 73*(8), 311-316.

Fore, C., Burke, M. D., & Martin, C. (2006). Curriculum-based measurement: An emerging alternative to traditional assessment for African American children and youth. *The Journal of Negro Education*, *75*, 16-24.

Good, R., & Jefferson, G. (1998). Contemporary perspectives on curriculum-based measurement validity. In Shinn, M. R., Editor, *Advanced applications of curriculum-based measurement.* New York, NY: Guildford Press.

Goodenow, C., & Grady, K. E. (1993). The relationship of school belonging and friends' values to academic motivation among urban adolescent students. *The Journal of Experimental Education, 62*(1), 60-71.

Griner, A. C., & Stewart, M. L. (2013). Addressing the achievement gap and disproportionality through the use of culturally responsive teaching practices. *Urban Education, 48* (4), 585-621.

Gunnarsdottir, T., Njardvik, U., Olafsdottir, A. S., Craighead, L. W., & Bjarnason, R. (2012). Teasing and social rejection among obese children enrolling in family-based behavioural treatment: effects on psychological adjustment and academic competencies. *International Journal of Obesity, 36*, 35–44.

Hodge, S. R., & Vigo-Valentin, A. (2014). *Health, nutrition, and physical activity.* In H. R. Milner IV and K. Lomotey (Eds.), Handbook of Urban Education. New York, NY: Routledge.

Katz, D. L. (2013). Childhood obesity trends in 2013: Mind, matter, and message. *Childhood Obesity, 9* (1), 1-2.

Keller-Margulis, M. A., Shapiro, E. S., & Hintze, J. M. (2008.) Long-term diagnostic accuracy of curriculum-based measures in reading and mathematics. *School Psychology Review, 37,* 374.

Kim, S., Mazza, J., Zwanziger, J., & Henry, D. (2014). School and behavioral outcomes among inner city children: Five-year follow-up. *Urban Education, 49* (7), 835-856.

Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics, 29* (1), 67-101.

Marston, D. B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. *Curriculum-based measurement: Assessing special children*, New York NY: Guilford Press.

McGlinchey, M. T., & Hixson, M. D. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review*, *33,* 193.

Milner, H. R. (2013). Analyzing poverty, learning, and teaching through a critical race theory lens. *Review of Research in Education*, *37* (1), 1-53.

Milner, H. R., & Lomotey, K. (2014). *Introduction.* In H.R. Milner & K. Lomotey (Eds.) Handbook of Urban Education. New York, NY: Routledge.

Naglieri, J. A., & Bornstein, B. T. (2003). Intelligence and achievement: Just how correlated are they? *Journal of Psychoeducational Assessment*, *21*, 244-260.

National Center on Student Progress Monitoring. What the National Center on Student Progress Monitoring can do for you 2007. Available from: http://www.studentprogress.org/library/Webinars/NCSPMWebinar8-30-07.pdf.

Sallis, J. F., McKenzie, T. L., Kolody, B., Lewis, M., Marshall, S., & Rosengard, P. (1999). Effects of health-related physical education on academic achievement: Project SPARK. *Research Quarterly for Exercise and Sport*, *70*(2), 127-134.

Sallis, J.F., Owen, N., & Fisher. E.B. (2008). Ecological Models of Health Behavior. In Glanz, K., Rimer, B.K., & Viswanath, K. (Eds.). *Health Behavior and Health Education: Theory, Research, and Practice* (4th ed., pp.465-485). San Francisco: Jossey-Bass.

Sameroff, A. (2010). A unified theory of development: A dialectic integration of nature and nurture. *Child Development, 81*(1), 6–22.

Sanders, W. L. (1998). Value-added assessment. *The School Administrator,* 55 (11), 24-32

Santoro, L. E., Lutz, J. G., Keller, M. A., Hintze, J. M., & Shapiro, E. S. (2006). Curriculum-based measures and performance on state assessment and standardized tests. *Journal of Psychoeducational Assessment, 24,* 19.

Shinn, M. R. (1998). Advanced applications of curriculum-based measurement. New York: Guilford Press.

Shinn, M. R, & Shinn, M. (2002). AIMSweb Training Workbook: Administration and Scoring of Reading Maze for Use in General Outcome Measurement. Eden Prairie MN: Edformation, Inc.; Available from: http://www.aimsweb.com/wp-content/uploads/Maze-Manual.pdf.

Shen, B. (2015). Physical education and academic performance in urban African American girls. *Urban Education*, DOI: 10.1177/0042085914566095, first published on January 4, 2015.

Singh, A., Uijtdewilligen, L., Twisk, J.R., van Mechelen, W., Chinapaw, M.M. (2012). Physical activity and performance at school: A systematic review of the literature including a methodological quality assessment. *Archives of Pediatric and Adolescent Medicine, 166*(1), 49-55. doi:10.1001/archpediatrics.2011.716.

Somers, C. L., & Gizzi, T. J. (2001). Predicting adolescents' risky behaviors: The influence of future orientation, school involvement, and school attachment. *Adolescent & Family Health, 2*(1), 3-11.

Stokols, D. (1992). Establishing and maintaining healthy environments: toward a social ecology of health promotion. *American Psychologist*, *47*(1), 6.

Steinberg, L. (2014). *Adolescence (10th ed.).* New York: McGraw-Hill.

Thurber, R. S., Shinn, M. R., & Smolkowski, K. (2002). What is measured in mathematics tests? Construct validity of curriculum-based mathematics measures. *School Psychology Review, 31*, 498.

Tindal, G., Helwig, R., & Anderson L. (2002). Using a concept-Grounded, Curriculum-Based Measure in Mathematics to Predict Statewide Test Scores for Middle School Students with LD. *The Journal of Special Education, 36,* 102.

Tolar, T. D, Vaughn, S, Stuebing, K. K., Barth, A. E., Fletcher, J. M., & Francis, D. J. (2012). Psychometric properties of maze tasks in middle school students. *Assessment for Effective Intervention, 37,* 131.

Trudeau, F., Shepard, R. J. (2008). Physical education, school physical activity, sports and academic performance. *International Journal of Behavioral Nutrition and Physical Activity,* 5 (10). doi:10.1186/1479-5868-5-10

Ullucci, K., & Howard, T. (2015). Pathologizing the poor: Implications for preparing teachers to work in high-poverty schools. *Urban Education, 50 (*2), 170-193.

Voelkl, K. E. (1996). Measuring students' identification with school. *Educational and Psychological Measurement, 56,* 760-770. doi: 10.1177/0013164496056005003

Wang, M. T., & Eccles, J. S. (2012b). Social support matters: Longitudinal effects of social support on three dimensions of school engagement from middle to high school. *Child Development*, *83*, 877-895. doi: 10.1111/j.1467- 8624.2012.01745.x.

Weber, N. M., Somers, C. L., Day, A., & Baroni, B. A. (2016). Predictors and outcomes of school attachment and school involvement in a sample of girls in residential treatment. *Residential Treatment for Children and Youth.* Published online first: http://www.tandfonline.com/doi/full/10.1080/0886571X.2016.1188034

Yeo, S. (2010). Predicting performance on state achievement tests using curriculum-based measurement in reading: A multilevel meta-analysis. *Remedial and Special Education, 31*, 412-422.

Table 1

Means and Standard Deviations for the Full Sample

| Variable | Pre-Test | | Post-Test | | Pair Samples T-test t; p |
|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | |
| Math scores | 25.67 | 16.80 | 42.55 | 19.04 | -19.66; .000 |
| Reading scores | 14.74 | 7.87 | 22.59 | 9.16 | -18.74; .000 |
| School attachment | 4.04 | .87 | 3.86 | .89 | 3.5; .001 |
| Peer acceptance | 4.27 | .86 | 4.29 | .80 | -.46; .648 |
| Math ROI | -- | -- | .56 | .48 | n/a |
| Reading ROI | -- | -- | .26 | .23 | n/a |

*Note.* ROI=Rate of Improvement

Table 2

Means, Standard Deviations, and Paired Samples t-tests by School

| Variable | Pre-Test | | Post-Test | | Paired Samples t-test t; df; p |
|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | |
| *School 1—19% FRPM public; 301 children* | | | | | |
| School attachment | 4.10 | .76 | 4.12 | .68 | ns |
| Peer acceptance | 4.30 | .73 | 4.41 | .58 | ns |
| Math scores | 28.98 | 14.96 | 54.64 | 14.6 | -16.49; 46; .000 |
| Reading scores | 19.09 | 8.18 | 28.62 | 8.80 | -13.56; 44; .000 |
| *School 2—51% FRPM; private; 327 children* | | | | | |
| School attachment | 3.91 | 1.05 | 3.57 | .88 | ns |
| Peer acceptance | 4.33 | 1.07 | 4.21 | .96 | ns |
| Math scores | 58.48 | 12.28 | 61.56 | 10.15 | ns |
| Reading scores | 20.74 | 6.16 | 23.89 | 7.52 | -3.21; 26; .004 |
| *School 3—77% FRPM; public; 655 children* | | | | | |
| School attachment | 4.03 | .76 | 3.67 | .95 | 3.25; 58; .002 |
| Peer acceptance | 4.21 | .86 | 4.31 | .81 | ns |
| Math scores | 23.16 | 12.46 | 37.75 | 17.50 | -9.23; 72; .000 |
| Reading scores | 12.96 | 7.43 | 24.24 | 9.96 | -12.43; 71; .000 |
| *School 4—71% FRPM; public; 310 children* | | | | | |
| School attachment | 4.55 | .46 | 4.34 | .75 | ns |
| Peer acceptance | 4.36 | .87 | 4.28 | .91 | ns |
| Math scores | 23.74 | 10.58 | 39.18 | 14.21 | -7.38; 37; .000 |
| Reading scores | 14.86 | 8.40 | 18.5 | 8.45 | -5.37; 57; .000 |
| *School 5—78% FRPM; public; 287 children* | | | | | |
| School attachment | 3.75 | .94 | 3.77 | .72 | ns |
| Peer acceptance | 4.30 | .70 | 4.21 | .71 | ns |
| Math scores | 17.36 | 10.25 | 44.31 | 17.77 | -12.62; 38; .000 |
| Reading scores | 13.42 | 5.60 | 18.24 | 6.96 | -7.06; 37; .000 |
| *School 6—81% FRPM; public; 699 children* | | | | | |
| School attachment | 3.84 | .78 | 3.82 | .68 | ns |
| Peer acceptance | 4.17 | .92 | 4.27 | .85 | ns |
| Math scores | 19.07 | 14.68 | 31.97 | 19.84 | -7.81; 59; .000 |
| Reading scores | 9.87 | 5.41 | 22.03 | 7.13 | -10.71; 37; .000 |

*Note.* FRPM=Free and Reduced Price Meals; also indicated is whether each school is public or private and the number of children in each school denotes a relatively smaller or larger school. ns=not significant.

Table 3

ANOVA Results: Math and Reading ROI by School

| Variable | Math ROI | | Reading ROI | |
|---|---|---|---|---|
| | Mean | SD | Mean | SD |
| *School 1—19% FRPM public; 301 children* | .86 $_c$ | .36 | .32 $_b$ | .16 |
| *School 2—51% FRPM; private;327children* | .10 $_a$ | .42 | .10 $_a$ | .17 |
| *School 3—77% FRPM; public; 655 children* | .49 $_b$ | .45 | .38 $_b$ | .26 |
| *School 4—71% FRPM; public; 310 children* | .51 $_b$ | .43 | .12 $_a$ | .17 |
| *School 5—78% FRPM; public; 287 children* | .89 $_c$ | .44 | .16 $_a$ | .14 |
| *School 6—81% FRPM; public; 699 children* | .43 $_b$ | .43 | .41 $_b$ | .23 |

*Note*. Matching subscripts denote mean scores that are *not* statistically significantly different than each other. The national 4$^{th}$ grade ROI is .72 for math and .40 for reading comprehension. FRPM=Free and Reduced Price Meals; also indicated is whether each school is public or private and the number of children in each school denotes a relatively smaller or larger school.