

# High/Scope Youth PQA Technical Report



## Afterschool Quality and School-Day Outcomes

**Juliane Blazevski and Charles Smith**

High/Scope Educational Research Foundation



## Summary

Findings from this study suggest that several dimensions of youth program quality are associated with a range of positive school-day behavioral and academic outcomes. Analyses suggest that the elements of program quality that influence youth program attendance may be different from the elements of program quality that are most predictive of positive youth outcomes. Opportunities for reflection on daily work were positively associated with standardized reading measures. Findings should be interpreted with caution since they are based on a relatively small sample, and the elements of research design employed in the study are not optimal.

## Introduction

The Youth Program Quality Assessment (PQA) is an observational assessment tool developed by the High/Scope Educational Research Foundation that describes the quality of developmental and learning experiences available to children and youth in afterschool programs. This paper is part of a series of reports that present evidence regarding the consistency and meaningfulness of data produced on the Youth Program Quality Assessment. This piece of the validation research speaks to issues of predictive validity — are Youth PQA data for afterschool programs associated with school-day outcomes for the children that attend afterschool programs in those same schools? This is a challenging question, in large part because afterschool and school-day settings for any individual child may occur in the same building but may be otherwise quite disconnected. The simple assumption that afterschool programs will automatically increase academic performance has been challenged by arguments based on both common sense (Kane, 2004; Halpern, 2005) and empirical data (James-Burdumy, 2005).

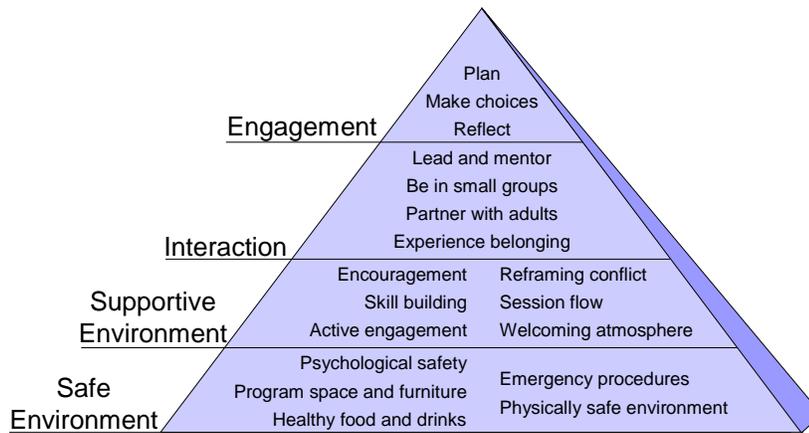
In earlier analyses, quality scores from the Youth PQA's interaction and engagement domains were related to youth reports of challenge, skill building, and sense of growth from participation in afterschool programs (Smith & Hohmann, 2005), and these findings have been replicated (Smith et al., 2007). At least two other recent evaluations that employ well-elaborated quality measures suggest that afterschool program quality matters for a wide range of socio-emotional and cognitive outcomes as well, including school performance (ICRE & NIOST, 2005; Russell & Reisner, 2005). Recent meta-analyses suggest that improved school performance is related to the quality of afterschool offerings (Durlak & Wiessberg, 2007) and that school-day effects are not limited to afterschool programs that focus strictly on academic work (Durlak & Wiessberg, 2007; Lauer, et al., 2000).

The policy context for this research is the federally funded Michigan Department of Education 21<sup>st</sup> Century Community Learning Centers program. The 21st Century policy explicitly states that the primary purpose of the program is increased academic achievement (McCallion, 2003), so seeking empirical association between afterschool experiences and school-day outcomes is mandated in this policy context. This study is focused on the quality of experiences available to children and youth in afterschool contexts, rather than on attendance or other traditional ways of estimating the dosage of afterschool programs. We believe it is likely that the quality of afterschool experiences — e.g., *nurturing relationships* during homework help, the experience of *collaboration with peers* on a life-skills lesson, *powerful reflection* on a community service project — may be catalysts of developmental change related to motivation, identity, and learning that find application during the school day.

## Background

The Youth PQA is an assessment of best practices in afterschool and community programs for youth. Procedures and methods for use of this observational assessment are described in detail elsewhere (High/Scope, 2005). The Youth PQA consists of 4 domains and 18 scales pictured in Figure 1. Items are scored from 1 to 5 and then averaged up to the scale (multi-item) and domain (multi-scale) levels. The Youth PQA demonstrates acceptable levels of reliability and validity based on widely shared psychometric criteria (Smith & Hohmann, 2005; Yohalem et al., 2007).

# Pyramid of Program Quality



1

**Figure 1. Program quality scores**

The pyramid shape represents an empirical trend present in nearly every recent study using the Youth PQA — scores are highest at the base (Safe Environment), getting gradually lower as the pyramid narrows (Smith, et al., 2006; Learning Points Associates & Berkley Policy Associates, 2006; Gramiak et al., 2006).

In the 2004–2005 school year, the Grand Rapids Public School District contracted with High/Scope to collect Youth PQA data in all 21 of its Michigan 21<sup>st</sup> Century Community Learning Centers, each located in a district school building and operated by one of three nonprofit management organizations in the city (City of Grand Rapids Department of Recreation, Boys and Girls Clubs, YMCA). More recently, the school district provided High/Scope with an individual student level data set for participants in the same afterschool sites that the district had already established to evaluate the after-school sites during 2004–2005 school year.

The circumstances of this “accidental” sample imply several caveats for our findings. First, our level 2 sample sizes are very small when we attempt to break the data file by school level. Second, while we were fortunate to receive child-outcome data from the previous school year to use as controls, we suspect that the models are underspecified. We lack important explanatory variables in our quantitative models, especially at the school-context level. Finally, the study is based on cross-sectional and correlational data. This suggests that our analytic framework — the elements of research design employed to support interpretation of the findings — is relatively weak.

Our aims for this research are primarily related to instrument development and support for the use of observational assessment. We hope to provide information and experience for others who are considering the use of observational-setting measures in studies with more powerful longitudinal and experimental designs.

## Sample and Measures

The sample for the study includes 48 observations of afterschool programs collected at 21 sites — 7 middle- and 14 elementary-level — during the 2004–2005 school year. At each site, minimum data collection included

one complete Youth PQA rating for an academic enrichment offering led by a Grand Rapids Public Schools' teacher, and one complete Youth PQA rating for a general enrichment offering led by a youth worker employed by one of the three site management organizations.

This report also contains data collected from 3,484 youth attending the 21 afterschool programs. The youth sample contains approximately equal numbers of males and females in grades K–8. The sample is diverse with 52% African American, 34% Hispanic, and 12% Caucasian youth, and 2% other. Eighty-seven percent of the youth qualified for the free lunch program, 17% received special education services, and 19% were bilingual. Sixteen percent of the youth had been suspended during the 2004–2005 school year. Only around 7% of the sample scored at or above 50 normal curve equivalent (NCE) on the Gates-MacGinitie Reading Test (GMRT) (Gates-MacGinitie, 1989).

Quality measures for the study included four scales from Form A of the Youth PQA:

- I-C. Appropriate emergency procedures and supplies are present.** 6 items; No alpha reported; this type of scale does not meet reliability model assumptions)
  - Written emergency procedures are posted in plain view.
  - At least one charged fire extinguisher is accessible and visible from the program space.
  - At least one complete first-aid kit is accessible and visible from the program space.
  - Other appropriate safety and emergency equipment (e.g., for water or vehicle safety, sports, or repairs) is available to the program offering as needed, can be located by staff, and is maintained in full-service condition.
  - All entrances to the indoor program space are supervised for security during program hours.
  - Access to outdoor program space is supervised during program hours.
- III-L. Youth have opportunities to develop a sense of belonging.** (4 items; alpha = .41)
  - Youth have structured opportunities to get to know each other.
  - Youth exhibit predominantly inclusive relationships with all in the program offering, including newcomers.
  - Youth strongly identify with the program offering.
  - The activities include structured opportunities to publicly acknowledge the achievements, work, or contributions of at least some youth
- IV-P. Youth have opportunities to set goals and make plans.** (2 items; alpha = .66)
  - Youth have multiple opportunities to make plans for projects and activities.
  - In the course of planning the projects or activities, 2 or more planning strategies are used.
- IV-R. Youth have opportunities to reflect.** (4 items; alpha = .81)
  - All youth are engaged in an intentional process of reflecting on what they are doing or have done.
  - All youth are given the opportunity to reflect on their activities in 2 or more ways.
  - In the course of the program offering, all youth have structured opportunities to make presentations to the whole group.
  - Staff initiate structured opportunities for youth to give feedback on the activities.

All data were collected by a trained outside observer using a training methodology shown to produce acceptable levels of inter-rater agreement in the Youth Program Quality Assessment Validation Study (Smith & Hohmann, 2005).

Afterschool program attendance measures included: the number of sessions attended for specific content offerings and the number of days a child attended the afterschool program at all. School-day academic outcomes included: NCE scores for the GMRT during the 2004–2005 school years (Gates-MacGinitie, 1989). School-day behavioral outcomes included short-term and long-term suspensions during the 2004–2005

school years. Individual level background data included ethnicity, free lunch status, special education status, prior year short- and long-term suspensions, and prior year NCE scores for the GMRT.

## Analysis Strategy and Theoretical Approach

The overall analysis strategy for this study is inductive and exploratory. In general, we started with setting quality scores and individual outcomes aggregated to the level of each setting — in all cases the level of the entire site ( $N = 21$ ). Then we looked for bivariate relationships that were theoretically meaningful, substantive in magnitude, and statistically significant. Finally, we tested these relationships in multilevel statistical models using as much of the available data as possible. Our primary goal in conducting these analyses is to advance understanding of how Youth PQA scores operate in fairly simple linear models in order to inform instrument choice decisions in future studies that employ more complete measures and stronger designs.

Although our approach to analyses was inductive, we did begin the process with a very basic theory which is described in Figure 4. Simply stated, we believed that the quality of afterschool programs might have both indirect and direct effects on youth school-day outcomes. Two aspects of the school district context demonstrate that the afterschool programs were tightly coupled with the school-day and provide support for the plausibility of our findings that afterschool quality can be linked to school behavior and achievement in a simple model. First, academic content in the Grand Rapids Public Schools' afterschool programs during 2004–2005 was well articulated with school-day academic content. The afterschool program hired certified teachers from the same school buildings in which the afterschool programs were located to deliver academic content for the first hour of the program each day. The district also invested heavily in high quality academic enrichment curricula and materials (e.g., Lego Math) and, according to our data, the teachers delivered these activities using best practices for instruction. Finally, the district and the afterschool program invested substantial effort to ensure that building principals and school-day teachers were linked to the afterschool program.

## Findings

### Overall levels of program quality for 21<sup>st</sup> Century Programs, Grand Rapids Public Schools, 2004–2005

Table 1 presents overall quality scores for the 21<sup>st</sup> Century sites operated in 21 of the Grand Rapids Public Schools during the 2004–2005 school year. Scores can be interpreted using the following rule: a score of 5 suggests that the staff practice is available frequently for all of the children and youth in attendance; a score of 3 suggests that the practice occurs in the setting but not for all children and youth in attendance; a score of 1 suggests that the practice does not occur or occurs only infrequently in the setting.

For each site a rating was captured for an academic enrichment offering delivered by a certified teacher in the Grand Rapids district (columns headed *academic* in Table 1) and a general enrichment offering delivered by a youth worker (columns headed *enrichment* in Table 1) employed through the three afterschool management organizations.

**Table 1. Overall Quality Scores for Grand Rapids Public Schools' 21<sup>st</sup> Century Sites, 2004–2005 School Year**

	Middle School N = 7 sites, 14 Ratings	Elementary School N = 14 Sites, 28 Ratings	Norm* Scores, N = 175 sites, 635 Ratings		
	<u>Middle School Enrich- ment</u>	<u>Middle School Aca- demic</u>	<u>Elemen- tary School Enrich- ment</u>	<u>Elemen- tary School Aca- demic</u>	
<b>Safe Environment</b>					
Promoting psychological and emotional safety	4.07	4.57	4.36	4.92	4.31
Safe and healthy physical environment for youth	4.68	4.79	4.39	4.88	4.70
Emergency procedures and supplies	3.62	3.45	4.05	4.15	3.97
Physical environment accommodates program offerings	4.70	4.93	4.50	4.54	4.67
<b>Supportive Environment</b>					
Providing a warm and caring atmosphere	4.02	4.61	3.57	4.69	4.23
Activities are planned, presented and paced for youth	4.14	4.50	4.49	4.35	4.34
Providing opportunities to be actively engaged	4.21	4.28	3.68	3.85	3.61
Supporting youth in developing skills	3.57	4.71	2.86	4.00	3.52
Encouraging youth through supportive strategies	3.62	4.14	3.19	3.41	3.40
Using youth-centered approaches to resolving conflicts	3.32	3.00	2.47	3.67	3.17
<b>Opportunities for Interaction</b>					
Helping to develop a sense of belonging	3.04	3.21	3.61	3.65	3.41
Providing opportunities to participate in small groups	2.57	2.90	2.71	2.23	2.41
Opportunities for developing leadership skills	2.19	2.62	2.19	2.18	2.61
Engaging youth as partners	3.64	4.00	2.71	3.62	3.50
<b>Engaged Learning</b>					
Providing opportunities to set goals and make plans	2.14	2.71	2.93	2.00	2.11
Providing opportunities to make choices based on interests	3.57	3.14	3.29	2.85	2.86
Providing opportunities to reflect	3.07	2.71	3.93	2.42	2.51
<b>Domains</b>					
Safe Environment	4.27	4.43	4.33	4.62	4.41
Supportive Environment	3.82	4.35	3.50	3.97	3.78
Opportunities for Interaction	2.86	3.18	2.81	2.92	2.96
Engaged Learning	2.92	2.86	3.38	2.42	2.53

\*The norm column does not represent an intentional or representative sample. Norm scores are the simple average scores for approximately 635 offerings within 170 programs collected during the years 2001–2006 from various samples.

In general, quality scores for the Grand Rapids system are at or above average levels of quality for all of the other Youth PQA ratings we have collected using similar methods in recent years (see Norms column).

Indeed, given that quality scores for elementary level programs are usually lower than for middle and high school programs, the Grand Rapids scores are quite high on average. Furthermore, the quality of performances by certified teachers delivering academic content reflected in the Table 1 are not dramatically different than those for professional youth workers employed by youth services agencies to deliver enrichment content.

### Relationship Between Program Quality and Attendance<sup>i</sup>

Zero-order correlations between program quality, program attendance, and youth outcomes are shown in Table 2 (aggregated child-level and program-level data). Only variables with at least one significant (or marginally significant) correlation coefficient are included. Relationships between program quality, program attendance, and youth outcomes are highlighted in grey.

**Table 2. Correlations Between Program Quality, Program Attendance and Youth Outcomes ( N = 21)**

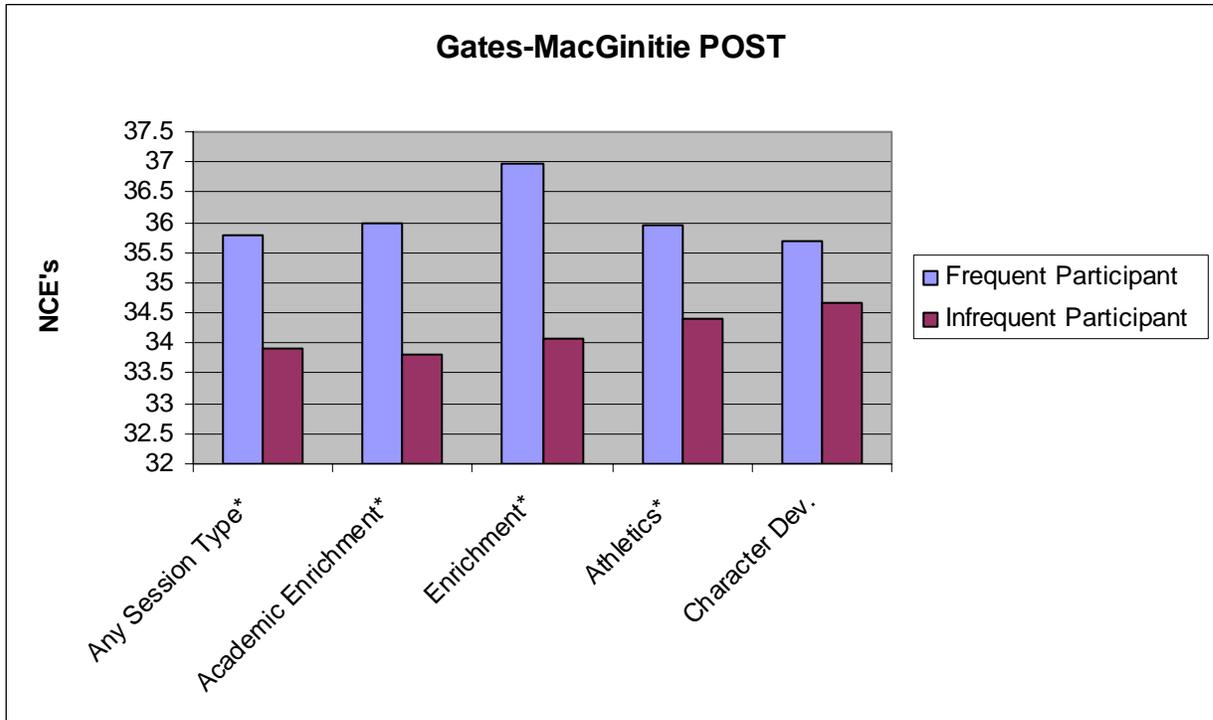
Variables	Reading (Gates) Post-Test	Reading (Gates) Gain Scores	Short-Term Suspension	Long-Term Suspension	Program Days Attended	Program Sessions Attended	Appropriate Emergency Procedures	Sense of Belonging	Set Goals / Make Plans	Reflect on Work
Reading (Gates) Post-Test	—									
Reading (Gates) Gain Scores	.602*	—								
Short-Term Suspension	-.288	-.398 ~	—							
Long-Term Suspension	.166	-.007	.299	—						
Program Days Attended	.371 ~	.358	-.484*	-.363	—					
Program Sessions Attended	.252	.588*	-.222	-.397 ~	.778*	—				
Appropriate Emergency Procedures	.117	.462*	-.262	-.328	.287	.421 ~	—			
Sense of Belonging	.395 ~	.299	-.173	-.306	.622*	.417 ~	.137	—		
Set Goals / Make Plans	-.118	.090	.040	-.516*	.092	.321	.402 ~	.145	—	
Reflect on Work	.428 ~	.455*	.011	-.060	.079	.267	.357	.244	.510*	—

As shown in Table 2, results suggest that the extent to which a program has *appropriate emergency procedures* in place is related to the average number of sessions/events that youth attended at a program ( $r = .42$ , marginally significant). Scores on another Youth PQA scale, *sense of belonging*, were positively correlated with the average number of days that youth attended a site ( $r = .62$ ). The relationship between *sense of belonging* and attendance was confirmed in more rigorous tests using multivariate and multilevel models when controlling for school type (elementary/middle), ethnicity, SES, and special education status (see Appendix: Table A)

### Relationship between Program Attendance and Youth Outcomes

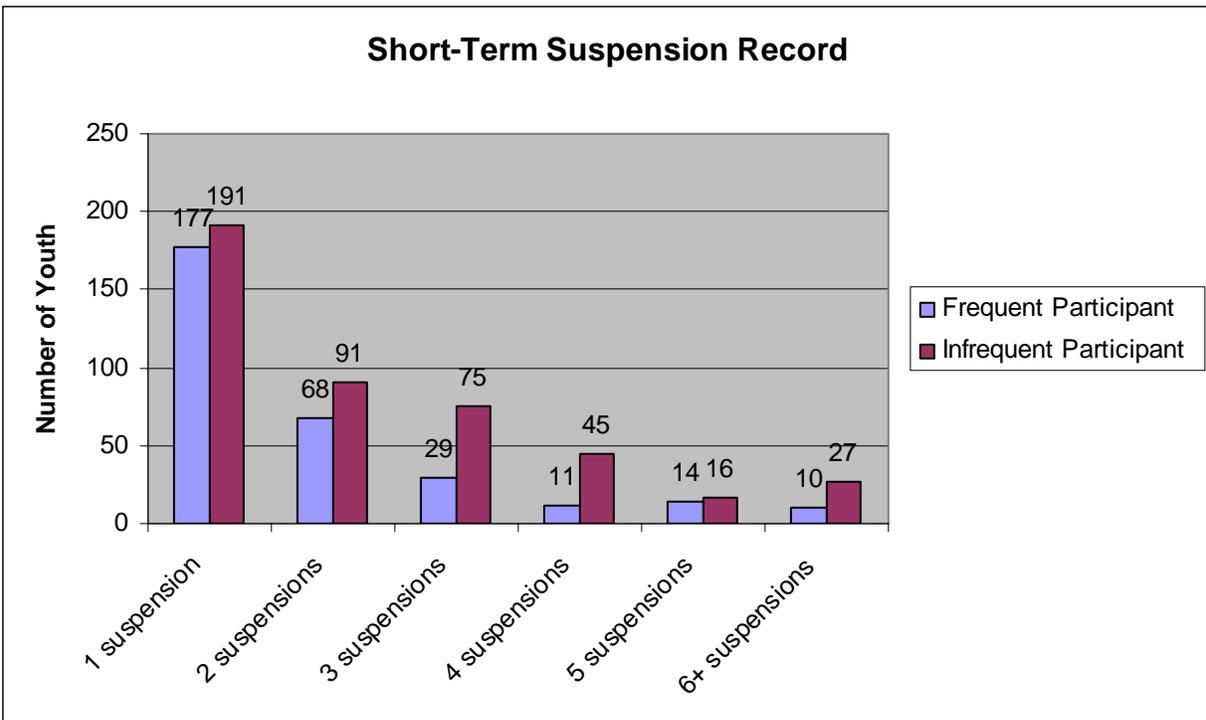
**Reading achievement.** Results suggest that attending a youth program may have benefits in terms of youth achievement in reading. As indicated in table 1, the number of days that youth attended a program (on average) is marginally correlated ( $r = .37$ ) with end-of-year reading scores on the Gates-MacGinitie assessment. An even stronger positive correlation ( $r = .59$ ) was identified between the average number of youth program sessions that youth attended and *gains* in reading scores (comparing Gates-McGinnitie pre- and post-test scores). Results of multivariate and multilevel models show some support of the positive relationship between program attendance and reading achievement. The number of sessions attended was a marginally significant predictor of Gates-MacGinitie 2005 scores when controlling for 2004 Gates-MacGinitie

scores, school type (elementary/middle), ethnicity, socioeconomic status (SES), and special education status (see Appendix: Table B). Figure 2 presents a graphic representation of differences in reading achievement by high and low levels of attendance at the afterschool program for any session type (see first set of bars) and for different types of program offerings (i.e., academic, enrichment, athletics, and character development).



**Figure 2**

**Suspension rates.** Program attendance also may have benefits in terms of improved youth behavior (as indicated by the number of suspensions reported). Specifically, the average number of days youth attended a program was negatively correlated with short-term suspensions ( $r = -.48$ ) and the average number of sessions attended was negatively correlated with long-term suspensions ( $r = -.40$ , marginally significant). Results of multivariate and multilevel models confirm the positive relationship between program attendance and short-term suspensions when controlling for school type (elementary/middle), and suspensions during the prior year (see Appendix: Table C). Figure 3 presents a graphic representation of differences in short-term suspension rates by high and low levels of attendance at the afterschool program.



**Figure 3**

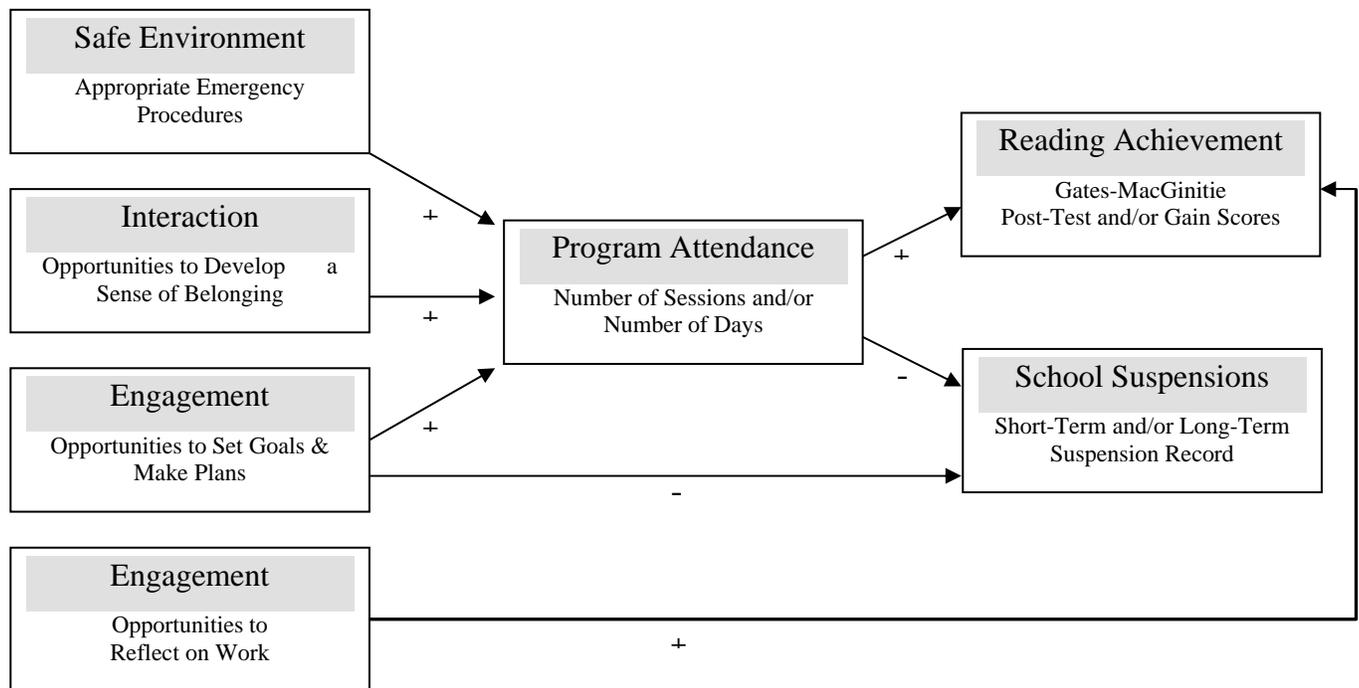
### Relationship between Program Quality and Youth Outcomes

**Reading achievement.** Analyses suggest that programs that provide opportunities to reflect on one's work may have a positive influence on youth's achievement in reading. Site scores on the Youth PQA scale *Youth have opportunities to reflect* (Form A, Scale IV.R) was positively correlated with mean Gates-MacGinitie post-test scores and gain scores ( $r = .43$  and  $.46$  respectively) for youth attending the program. Results from multivariate and multilevel tests of this relationship indicate that while much of a student's reading achievement can be explained by child-level variables, such as a youth's prior reading ability, special education status, ethnicity and SES, significant differences in youths' 2005 Gates-MacGinitie scores are also accounted for by a site's score on the scale *Youth have opportunities to reflect* (see Appendix: Table B).

**Suspension rates.** Analyses indicated that scores on the Youth PQA scale *Youth have opportunities to set goals and make plans* (Form A, Scale IV.P) are negatively associated with suspensions. That is, the higher a site's score on this scale the lower the average number of long-term suspensions (on average) for youth attending that program ( $r = -.52$ ; see Table 1). Multivariate and multilevel tests confirmed this relationship. After controlling for significant individual differences (prior suspension record, school type (elementary/middle), and youth program attendance), a site's score on the Youth PQA scale *Youth have opportunities to set goals and make plans* was a significant negative predictor of long-term suspensions (see Appendix: Table C).

### General Model for Afterschool Quality and School-Day Outcomes

Below is a model that depicts the general relationship between program quality, youth program attendance, and youth outcomes. Although this report highlighted relationships between *specific* Youth PQA scales and *specific* indicators of program attendance, reading achievement, and school suspension, this model is intended to tie results back to the general framework of youth program quality as conceptualized in Figure 1.



**Figure 4. General model of the relationship between program quality, program attendance, and youth outcomes.**

*Note:* This model depicts relationships suggested by a series of analyses (correlations, regression, HLM). This is not a true structural equation model. It is intended to be used as a way of summarizing the findings presented in this report and as a stimulus for further discussion.

## Discussion

For both researchers and policymakers, this study increases an understanding of relationships between afterschool program quality and students' school-day outcomes. For researchers, our findings support investment in studies that incorporate observational measures of program quality within more expensive research designs that include longitudinal follow-up and contrast conditions. A critical area for future study suggested by our research concerns differential relationships between afterschool quality and school-day outcomes by age of the student.<sup>ii</sup>

For public policy, the most important message that should be gleaned from this report is that afterschool program quality can make a difference in school-day outcomes. Results suggest that attendance rates are higher at quality programs that support youths' sense of belonging. Results also suggest that youth who frequently attend program sessions are likely to reap benefits in terms of their achievement and that providing youth with opportunities to reflect on their work may be key support in this equation. Finally, attending a youth program, particularly one in which there are opportunities to take ownership of program routines and purposes through frequent practices of goal setting and planning, may decrease the likelihood that youth end up suspended from school.

## References

- Cohen, J. (1988). *Statistical power for the behavioral sciences* (2<sup>nd</sup> Ed.). Hillsdale, NJ: Erlbaum.
- Durlak, J. A., & Weisberg, R. P. (2007). *The impact of afterschool programs that promote personal and social skills*. Chicago: Collaborative for Academic, Social, and Emotional Learning.
- Gates-MacGinitie Reading Tests (3<sup>rd</sup> Ed.). (1989). Rolling Meadows, IL: Riverside..
- Gramiak, W., Vanauken, L., Brugger, L., & Young-Miller, G. M. (2006). *Greater Rochester afterschool alliance (GRASA) assessment*. Rochester, NY: Children's Institute.
- Halpern, D. (2005). *Confronting the big lie: The need to reframe expectations of afterschool programs*. Chicago: Erickson Institute.
- Intercultural Center for Research in Education, & National Institute on Out-of-School Time. (2005). *Pathways to success for youth: What counts in afterschool*. Arlington, MA: United Way of Massachusetts Bay.
- High/Scope Educational Research Foundation. (2005). *Youth Program Quality Assessment*. Ypsilanti, MI: High/Scope Press.
- James-Burdumy, S., Dynarski, M., Moore, M., Deke, J., Mansfield, W., Pistorino, C., & Warner, E. (2005). *When schools stay open late: The national evaluation of the 21st Century Community Learning Centers Program final report*. Washington, DC: U.S. Department of Education, National Center for Education, Evaluation, and Regional Assistance.
- Kane, T. J. (2004). *The impact of afterschool programs: Interpreting the results of four recent evaluations*. New York: W. T. Grant Foundation.
- Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. A., Snow, D., & Martin-Glen, M. (2003). *The effectiveness of out-of-school time strategies in assisting low-achieving students in reading and mathematics*. Aurora, CO: Mid-continent Research for Education and Learning (McREL).
- Learning Point Associates & Berkeley Policy Associates. (2006). *South Carolina extended learning time study: Final report*. Columbia, SC: South Carolina Oversight Committee.
- McCallion, G. (2003). 21st Century Community Learning Centers in P.L. 107-110: Background and Funding (Order Code RL31240). Congressional Research Service Reports. Retrieved October 11, 2007, from University of North Texas Libraries Web site: <http://digital.library.unt.edu/govdocs/crs/index.tkl>
- Russell, C. A., & Reisner, E. R. (2005). *Supporting social and cognitive growth among disadvantaged middle-grades students in TASC afterschool project*. New York: The After-School Corporation.
- Smith, C., & Hohmann, C. (2005). *The Youth Program Quality Assessment validation study: Findings for instrument validation*. Ypsilanti, MI: High/Scope Educational Research Foundation.
- Smith, C., Akiva, T., & Henry, B. (2006). *Quality in the out-of-school time sector: Insights from the Youth PQA Validation Study*. Paper presented at the biennial meeting of the Society for Research on Adolescence, San Francisco, CA.
- Yohalem, N., & Wilson-Ahlstrom, A. (2007). *Measuring youth program quality: A guide to assessment tools*. Washington, DC: Forum for Youth Investment.

## Appendix

**Table A1**

*Results of HLM Explaining Variance in the Number of Days Youth Attended Afterschool Programs.*

*N = 21 afterschool programs; 1829 youth*

<u>Estimated Effects</u>	<u>Coefficient</u>
Intercept ( $\beta_0$ )	44.81**
Program-Level Variables:	
Support for Sense of Belonging	9.73*
Youth-Level Variables:	
School Type ( $\beta_1$ )	23.54**
Ethnicity ( $\beta_2$ )	-11.40**
SES ( $\beta_3$ )	-1.99
Special Ed ( $\beta_4$ )	-8.56**
<u>Inter-class coefficient (ICC)</u>	.163
<u>Proportion of within-program variance explained (youth-level)</u>	.011
<u>Proportion of between-program variance explained</u>	.299
<u>Lambda reliability estimate</u>	.755

Notes: \*  $p < .05$ ; \*\*  $p < .01$

School Type coded 0 = middle school, 1 = elementary school; Ethnicity coded 0 = not Caucasian, 1 = Caucasian; SES coded 0 = not qualified for free lunch status; 1 = qualified for free lunch status; Special Ed coded 0 = not qualified for special education services, 1 = qualified for special education services.

Coefficients are reported in their original metric. Intercept represents mean number of days attended; other coefficients represent changes in days predicted by either membership in a particular category (dummy coded score of "1"), or in the case of Support of Sense of Belonging, the coefficient represents the change in days attended associated with a unit increase above the mean for that variable (e.g., variable is centered around the grand mean).

**Table A2**

*Results of HLM Explaining Variance in Reading Achievement (Normal Curve Equivalent Scores on Gates-MacGinitie Reading Test, 2005).*

*N* = 21 afterschool programs; 1829 youth

<u>Estimated Effects</u>	<u>Coefficient</u>
Intercept ( $\beta_0$ )	36.64**
Program-Level Variables:	
Opportunities to Reflect on Work	1.01*
Youth-Level Variables:	
School Type ( $\beta_1$ )	2.24**
Ethnicity ( $\beta_2$ )	2.79**
SES ( $\beta_3$ )	-2.53**
Special Ed ( $\beta_4$ )	-4.82**
Gates-MacGinitie NCE 2004 ( $\beta_5$ )	.78**
Number of Sessions Attended ( $\beta_6$ )	.004~
Inter-class coefficient (ICC)	.051
<u>Proportion of within-program variance explained (youth-level)</u>	.596
<u>Proportion of between-program variance explained</u>	.218
<u>Lambda reliability estimate</u>	.406

Notes: ~  $p < .10$ , \*  $p < .05$ ; \*\*  $p < .01$ .

School Type coded 0 = middle school, 1 = elementary school; Ethnicity coded 0 = not Caucasian, 1 = Caucasian; SES coded 0 = not qualified for free lunch status, 1 = qualified for free lunch status; Special Ed coded 0 = not qualified for special education services, 1 = qualified for special education services.

Coefficients are reported in their original metric. Intercept represents mean NCE score on the Gates-MacGinitie 2005 test (i.e., standardized percentile rank) ; other coefficients represent changes in NCE predicted by either membership a particular category (dummy coded score of "1"), or in the case of Opportunities to Reflect on Work, Gates-MacGinitie NCE 2004, and Number of Sessions Attended, the coefficient represents the change in NCE associated with a unit increase above the mean for that variable (e.g., variables are centered around the grand mean).

**Table A3**

*Results of HLM Explaining Variance in the Number of Short-term and Long-term Suspensions During 2004-2005 School Year.*

*N* = 21 afterschool programs; 1829 youth

<u>Estimated Effects</u>	<u>Short-term Suspensions</u>	<u>Long-term Suspensions</u>
	Coefficient	Coefficient
Intercept ( $\beta_0$ )	.691**	.030*
Program-Level Variables:		
Opportunities to Set Goals & Make Plans	.009	-.022*
Youth-Level Variables:		
School Type ( $\beta_1$ )	-.361**	-.015
Short-term Suspension 2003-2004 ( $\beta_2$ )	.419**	.015**
Long-term Suspension 2003-2004 ( $\beta_3$ )	.560**	.221**
Number of Program Days Attended ( $\beta_4$ )	-.002**	-.000
<u>Inter-class coefficient (ICC)</u>	.110	.054
<u>Proportion of within-program variance explained (youth-level)</u>	.184	.070
<u>Proportion of between-program variance explained</u>	.000	.234
<u>Lambda reliability estimate</u>	.725	.706

Notes: \*  $p < .05$ ; \*\*  $p < .01$

School Type coded 0 = middle school, 1 = elementary school

Coefficients are reported in their original metric. Intercepts represent mean number of short-term and long-term suspensions; School Type coefficient represents predicted change in mean number of suspensions associated with being in an elementary (versus middle) school.

Coefficients for the other variables represent the predicted change in number of suspensions associated with a unit increase above the mean for that variable (e.g., variables are centered around the grand mean).

---

<sup>i</sup> The relationship between attendance and quality is complicated. First, in other samples also drawn from Michigan's 21<sup>st</sup> Century program, the relationship between attendance and program quality has been non-significant or, frequently, negative. As our colleague Ajay Kashew at The Afterschool Corporation has pointed out, the relationship between attendance and quality can easily be negative because it is easier to carry out some high-quality practices with smaller groups, so sites with lower attendance often achieve higher levels of quality at the point of service. In this study, we used the actual number of days attended at the level of the individual child as our attendance measure, rather than the average number of children in attendance, so our question was: does quality keep individual kids coming back?

<sup>ii</sup> We conducted additional analyses comparing elementary school and middle school youth which revealed a number of theoretically interesting findings. Because these analyses were based on very small samples — only 7 middle schools and 14 elementary school when breaking out by age — we do not include the full results in the body of the report but limit our presentation to discussion here.

First, results suggested that while *sense of belonging* may play a role in determining the attendance of youth of all ages, additional factors that are aligned with developmentally appropriate practices for adolescents appear to influence the attendance of middle school youth in afterschool programs. The Youth PQA scales *activities support active engagement*, *youth have opportunities to set goals and make plans*, *youth have opportunities to make choices based on their interests* were also related to attendance for middle school-aged youth. Second, results suggest the relationship between program attendance and reading achievement may be stronger at the elementary school level. Finally, there is some evidence that while program attendance (and a focus on *setting goals and making plans* within these programs) is associated with lower rates of suspension for youth of all ages, the extent to which programs provide *activities that support active engagement* and *opportunities to make choices* may also relate to lower suspension rates for middle school students.