

**Campbell Systematic Reviews**

2006:3

First published: 15 February, 2006

Last updated: 15 February, 2006

# **Impacts of After-School Programs on Student Outcomes**

Susan Goerlich Zief, Sherri Lauver, Rebecca A. Maynard



**THE CAMPBELL COLLABORATION**

# Colophon

---

<b>Title</b>	Impacts of after-school programs on student outcomes
<b>Institution</b>	The Campbell Collaboration
<b>Authors</b>	Goerlich Zief, Susan, Ph.D. Lauver, Sheri, Ph.D. Maynard, Rebecca A., Ph.D.
<b>DOI</b>	10.4073/csr.2006.3
<b>No. of pages</b>	53
<b>Last updated</b>	15 February, 2006

---

<b>Citation</b>	Zief SG, Lauver S, Maynard RA. Impacts of after-school programs on student outcomes. Campbell Systematic Reviews 2006:3 DOI: 10.4073/csr.2006.3
<b>Copyright</b>	© Zief et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
<b>Keywords</b>	

---

<b>Contributions</b>	None stated.
<b>Support/Funding</b>	Smith Richardson Foundation Hewlett Foundation
<b>Potential Conflicts of Interest</b>	Zief and Maynard coded this study in order to avoid a conflict of interest and determined that this study met the inclusion criteria for this review.

---

<b>Corresponding author</b>	Susan Goerlich Zief Mathematica Policy Research, Inc. 600 Alexander Park Princeton, NJ 08540 USA Telephone: +1 609 275 2291 E-mail: <a href="mailto:szief@mathematica-mpr.com">szief@mathematica-mpr.com</a>
-----------------------------	--

# Campbell Systematic Reviews

---

**Editors-in-Chief** Mark W. Lipsey, Vanderbilt University, USA  
Arild Bjørndal, Norwegian Knowledge Centre for the Health Services &  
University of Oslo, Norway

## **Editors**

*Crime and Justice* David B. Wilson, George Mason University, USA

*Education* Chad Nye, University of Central Florida, USA  
Ralf Schlosser, Northeastern University, USA

*Social Welfare* Julia Littell, Bryn Mawr College, USA  
Geraldine Macdonald, Queen's University, UK & Cochrane Developmental,  
Psychosocial and Learning Problems Group

**Managing Editor** Karianne Thune Hammerstrøm, The Campbell Collaboration

## **Editorial Board**

*Crime and Justice* David Weisburd, Hebrew University, Israel & George Mason University, USA  
Peter Grabosky, Australian National University, Australia

*Education* Carole Torgerson, University of York, UK

*Social Welfare* Aron Shlonsky, University of Toronto, Canada

*Methods* Therese Pigott, Loyola University, USA  
Peter Tugwell, University of Ottawa, Canada

---

The Campbell Collaboration (C2) was founded on the principle that systematic reviews on the effects of interventions will inform and help improve policy and services. C2 offers editorial and methodological support to review authors throughout the process of producing a systematic review. A number of C2's editors, librarians, methodologists and external peer-reviewers contribute.

The Campbell Collaboration  
P.O. Box 7004 St. Olavs plass  
0130 Oslo, Norway  
[www.campbellcollaboration.org](http://www.campbellcollaboration.org)

**Impacts of After-School Programs on Student Outcomes:  
A Systematic Review for the Campbell Collaboration**

**Susan Goerlich Zief, Ph.D.<sup>1,2</sup>**

Researcher  
Mathematica Policy Research, Inc.  
600 Alexander Park  
Princeton, NJ 08540  
609-275-2291  
szief@mathematica-mpr.com

**Sherri Lauver, Ph.D.<sup>1, 2</sup>**

Assistant Professor  
Warner Graduate School of Education and Human Development  
University of Rochester  
Dewey Hall 1-206A  
Rochester, NY 14627  
585-275-0057  
slauver@its.rochester.edu

**Rebecca A. Maynard, Ph.D.<sup>2</sup>**

Advisor to Reviewers  
University Trustee Professor of Education and Social Policy  
Graduate School of Education  
University of Pennsylvania  
3700 Walnut Street  
Philadelphia, PA 19104  
215-898-3558  
rmaynard@gse.upenn.edu

**Sources of Support**

Smith Richardson Foundation  
Hewlett Foundation

**Submitted**

February 2006

---

<sup>1</sup>Dr. Zief and Dr. Lauver completed this work in a prior position with the University of Pennsylvania, consulting to the Test-Bed Review Project for the Campbell Collaboration.

<sup>2</sup>The findings and views in the paper are those of the authors, and are neither a product of nor attributable to Mathematica Policy Research, Inc., the University of Rochester, nor the University of Pennsylvania.

## **Impacts of After-School Programs on Student Outcomes: A Systematic Review for the Campbell Collaboration**

### **Background for the Review**

In the last decade, the amount of after-school programming in the United States has grown tremendously. The number of schools that offered after-school programs doubled between 1994 and 2000 (De Kanter, 2001), at which time two-thirds of school principals reported that their schools offered these programs (Belden Russonello & Stewart Research and Communications, 2001). It is estimated that at least 6 million, or 11 percent, of school-age youth regularly participated in after-school programs during the 2002–03 school year (Afterschool Alliance, 2004). The federal government, states, localities, and private foundations have invested substantial resources in programs. For example, appropriations for 21<sup>st</sup> Century Community Learning Centers (CCLCs) increased from \$40 million in 1998 to the nearly \$1 billion appropriated for the program in 2004.

Increased pressure from parents and the public to bridge the gap between the end of the school day and the end of the work day has contributed to the great expansion in programming. It has recently been estimated that three million children in the United States between the ages of 6 and 12, or 15 percent of that age range, are frequently unsupervised after school (Vandivere, Tout, Zaslow, Calkins, & Capizzano, 2003). These numbers have increased as more women have entered the work force: over 80 percent of mothers of school-age youth now work outside of the home (Proscio & Whiting, 2004).<sup>3</sup> Additionally, more than two-thirds of low- and moderate-income youth do not have parental supervision available after school because of parental work requirements (Long & Clark, 1998; U.S. Bureau of Labor Statistics, 2000).

A second factor contributing to the recent growth in after-school programming is increased public recognition that unstructured, unsupervised after-school time presents “risk and opportunity” (Hofferth, 1995). Research has linked unsupervised time with increased risk-taking behaviors, victimization, and poor academic outcomes (Dwyer et al., 1990; Newman, Fox, Flynn, & Christeson, 2000; Osofsky, 1999; Posner & Vandell, 1999; Richardson et al., 1989; U.S. DHHS, 1995; U.S. Department of Education & U.S. Department of Justice, 2000). After-school programs have been touted as a means to reduce such negative outcomes and promote academic, social and emotional, and behavioral growth by providing positive supervision while also offering academic programming and recreational and youth development activities.

Some research suggests that after-school programs can result in improved outcomes for youth. However, this evidence is limited mainly to quasi-experimental and non-experimental studies that have correlated program participation with decreased participation in criminal or violent activities and improved developmental, social and emotional, and academic outcomes (Baker & Witt, 1996; Foley, Eddins, & Fenton, 2000; Grossman et al., 2002; Huang, Gribbons, Kim, Lee, & Baker, 2000; Jones & Offord, 1989; Le & Hamilton, 2001; McLaughlin & Irby, 1994; Posner & Vandell, 1994; Ross, Saavedra, Shur, Winter, & Felner, 1992; Schinke, Cole, & Poulin, 2000;

---

<sup>3</sup>In 1970, only 39 percent of mothers with children 18 years old and younger worked outside of the home (Proscio & Whiting, 2004).

U.S. Department of Education & U.S. Department of Justice, 2000; Welsh, Russell, Williams, Reisner, & White, 2002). Only recently have experimental design studies examined the impacts of participation in after-school programs on youth outcomes. And it is this type of study that offers the only reliable source of causal evidence on program impacts.<sup>4</sup>

This review examines the evidence from recent experimental design evaluations on the impact of after-school programming on youth context (i.e., student location, supervision, and safety); participation in activities; and behavioral, social and emotional, and academic outcomes. This review focuses on program models that are of particular interest to policymakers—programs that include academic support services—for two main reasons. First, 21<sup>st</sup> CCLC grantees are expected to incorporate academic support services into their programming, and it is likely that a large percentage of after-school programs are now funded with 21<sup>st</sup> CCLC grants. Second, as part of the No Child Left Behind legislation, the federal government legislated Supplemental Educational Services (SES), including after-school programs, to assist students who attend Title I schools not meeting performance goals.

### Recent Reviews of Evidence

There have been six major reviews of research on the impact of out-of-school programs that have included programs with an academic focus (Eccles & Templeton, 2002; Fashola, 1998; Hollister, 2003; Kane, 2004; Little & Harris, 2003; Scott-Little, Hamann, & Jurs, 2002).<sup>5</sup> However, all differ in important ways from this new review. For example, prior reviewers cast a broad net when defining an after-school program and included programs with a primary focus on tutoring, mentoring, youth development, or comprehensive services. Notably, most did not attempt a thorough or systematic search for all evidence, and one review limited inclusion to programs with “evidence of effectiveness” or programs that have “active dissemination and replicability materials” (Fashola, 1998, p. 5, 6). All reviews except Hollister (2003) included a mix of experimental, quasi-experimental, and even non-experimental studies.

These reviews generally reported some evidence of positive program impacts on academic and nonacademic outcomes (see Table 26 near the end of this document). However, many reviews did not describe their methods for assessing the quality of the studies on which they were reporting and did not address the fact that the evidence from which review conclusions were drawn varied considerably in terms of the internal validity of the estimates. Program design and delivery of the included studies was fairly heterogeneous, limiting the ability to draw conclusions across studies. Furthermore, the analyses were fairly descriptive, and no review attempted to combine the impact estimates from multiple studies using meta-analytic techniques.

---

<sup>4</sup>There is a strong body of research documenting the unreliability, and in some cases bias, in studies based on quasi-experimental studies, especially for studies of voluntary participation in programs (Guyat, DiCenso, Farewell, Willan, & Griffith, 2000; Agodini & Dynarski, 2001; Weisburd, Lum, & Petrosino, 2001; Glazerman, Levy, & Myers, 2003).

<sup>5</sup>Table 19 compares the methods and conclusions from prior reviews.

## Contribution of this Review

This review differs from prior reviews in four main ways. First, this review makes use of recently released experimental studies in the field that were not captured by all prior reviews. Second, this review focuses on a particular type of after-school program—those that operate on a regular basis after school during the school year and include some academic support services. Third, the review is based on a well-defined, systematic strategy for identifying all possible studies and determining those that are appropriate for inclusion in the review. All identified studies that have a relevant focus and that were judged to present credible evidence of program impacts were included, regardless of their findings. Fourth, this review pools the evidence of program impacts on particular outcomes across studies using meta-analytic methods.

## **Objectives of the Review**

This review was guided by three main questions:

1. To what extent and in what ways does access to after-school programs impact student context (i.e., student location, supervision, and safety), participation in enriching activities, behaviors, social and emotional development, and academic outcomes for youth?<sup>6</sup>
2. Do the effects of after-school programs vary among subgroups of youth defined by their baseline characteristics?
3. Among the program models and settings evaluated, do some seem more beneficial to youth than others? What are the distinguishing characteristics of those more and less successful programs?

This review includes four distinct components as it presents and discusses the available evidence that met predetermined standards for study quality:

1. Descriptions of the study designs, analytic methods, and program models
2. Tabular presentation of all measured outcomes, including indications of the direction and statistical significance of program impacts
3. Pooled estimates of the effect sizes and the probabilities that the overall effects are different from zero (where applicable)
4. A discussion of the findings and recommendations for strengthening the evidence base

---

<sup>6</sup>This review had initially planned to analyze behavioral, social and emotional, and academic outcomes of experimental design evaluations of relevant interventions. A full coding of all included studies revealed that most had also measured changes in student context (i.e., supervision, safety, parental involvement) and participation in activities. As these outcomes both have policy relevance and can theoretically be linked to changes in other outcome areas, we include them in this final version of the review.

## Review Methodology

The following presents a summary of the review methodology, which is detailed in a protocol approved by the Campbell Collaboration Education Coordinating Group.<sup>7</sup> The review procedures include criteria for study inclusion, the approach to identifying relevant studies and for judging whether they meet the study inclusion criteria, the content and process of coding data from the source studies, and the methods for analyzing and reporting the findings of the review.

### Study Inclusion Criteria

The review is limited to studies that focus on specific types of after-school programming, that serve relevant target populations and that measure at least one outcome of interest.

***Intervention and setting.*** This review is focused on after-school programs that combine recreation and/or youth development programming with academic support services. Mentoring, tutoring, summer, or in-school components could not be the primary means through which the programs attempt to influence outcomes. While mentoring and tutoring programs do operate during the after-school hours, the design and delivery of such programs assume a much different relationship between program teachers and youth, and therefore these programs fall out of the bounds of this particular review. Programs could operate in a variety of settings—schools, community centers, and religious institutions.

The search was limited to interventions implemented in North America. We did not specifically search for any international studies because the intervention context would likely be much different than that of U.S. and Canadian programs. Furthermore, Lauver's (2002) preliminary but extensive prior search found no relevant studies from countries other than the United States.

***Characteristics of the target population.*** Programs must serve youth enrolled in regular public or private K–12 schools; typically, these youth range in age from 5–19. The programs must not be targeted specifically at youth with special needs such as learning disabilities, physical disabilities, emotional problems, or behavioral problems.

***Types of studies.*** This review only included well-implemented experimental design studies because there is a strong body of research documenting the unreliability, and in some cases bias, in studies based on quasi-experimental studies (Guyat et al., 2000; Agodini & Dynarski, 2001; Weisburd, Lum, & Petrosino, 2001; Glazerman, Levy, & Myers, 2003). Furthermore, the study reports needed to provide a reasonable description of the methodology, program goals, and program activities.

The review included only studies published after 1982, 20 years prior to the initial search, for two primary reasons. First, bounding the search by 20 years is a developing standard for researchers conducting reviews for the Campbell Collaboration and the What Works Clearinghouse. Second, there was little public support for programs and evaluations in the field of after-school programming before the 1990s.

---

<sup>7</sup>A complete version of the approved protocol can be found at <http://www.campbellcollaboration.org/doc-pdf/afschprt.pdf>.

**Outcome measures.** The review initially sought to analyze three types of outcomes: (1) behavioral, (2) social and emotional, and (3) academic. Studies meeting all other inclusion criteria that reported findings for one or more of these outcomes were retained for review. The decision to include a study if it measured one or more of these outcomes is consistent with the fact that many logic models for after-school programs include a variety of outcomes, recognizing that many are interrelated (Appendix 1), and that non-experimental evidence shows after-school programs may have important impacts beyond those that map to their central goals (Eccles & Templeton, 2002; Grossman et al., 2002). Insofar as the included studies measured changes in student context and participation in activities, these outcomes are also reported.

### Literature Search Strategy and Results

This review employed a broad and systematic search for all studies meeting the specified inclusion criteria. This strategy included searching electronic databases, reference lists of prior reviews, the internet, and professional networks. The search terms were intentionally broad so that no relevant experimental work was missed. The search strategy did not include hand searching of journals, since prior experience indicated that it was highly unlikely that any rigorous experimental study would be missed through the other search methods.

These various search methods yielded five studies for inclusion in the review, none of which had been published in a peer-reviewed journal. The process of identifying these five studies involved scanning almost 1,000 citations and then retrieving and reviewing 88 studies (Table 1).

**Electronic databases.** We searched five major databases for this review—ERIC, Education Index, PsychINFO, C2 Spectr, and Dissertation Abstracts (Table 2).<sup>8</sup> The database search generated a total of 483 citations which were scanned for possible retrieval; 40 studies were retrieved and reviewed. Two experimental evaluations were identified (Tucker et al., 1998; Washburn, 2004) but were not included in the review because the program models were not relevant to this review (they were tutoring and obesity programs). A complete documentation of the database search strategy is included in Appendix 2.

**Prior reviews.** Through the six prior reviews cited above, as well as several less formal published collections of program names and evaluation descriptions, 110 separate potentially relevant studies were identified. Of these, only six were determined to be experimental design studies; three of the six met the criteria for the type of program to be included in this review (Table 3). The reviews that we searched for reference to experimental studies of after-school programs are noted in the reference section with one asterisk (\*).

---

<sup>8</sup>We did not search Sociological Abstracts because Lauver's preliminary and exhaustive search did not identify any relevant citations in this database.

Table 1: Summary of Search Results

Search source	Citations reviewed	Unduplicated studies retrieved and reviewed	Included studies
Electronic Databases			
ERIC	184	15	0
Education Index	21	0	0
PsychINFO	144	15	0
Dissertation Abstracts	125	9	0
C2 Spectr	9	1	0
Prior Reviews	110	46	3
Professional Networks	2	2	2
Internet			
google.com	400 of 2,936,000	0	0
Total	995	88	5

Table 2: Electronic Database Searches

Database	Search date	Search terms	Other parameters
Dissertation Abstracts	August 6, 2002	after school center, after school	
ERIC	August 28, 2002	program, after school education, after school AND evaluation, outcome, impact	1982–2002 English language
Education Index	August 29, 2002		
PsychINFO	October 2, 2002	after school center, after school program, after school <sup>9</sup>	None
C2 Spectr <sup>10</sup>	February 7, 2006	after school	Indexed or non-indexed fields

**Professional networks.** One study included in this review was conducted by one of the review authors (Lauver, 2002).<sup>11</sup> Another, the national study of 21<sup>st</sup> CCLCs, was identified and monitored through colleagues at Mathematica Policy Research, Inc. (U.S. Department of Education, 2003). The authors also used their contacts at the Harvard Family Research Project (HFRP) to identify any relevant studies that had not yet been documented in HFRP’s Out-of-School Time Evaluation Database.

<sup>9</sup>We used the same keywords for the ERIC, Education Index, and Dissertation Abstract searches. We were not able to use the same search terms in PsychINFO because the database identified “after” as a stop term, so the search results were highly irrelevant. Instead, a University of Pennsylvania librarian asked the management organization for the PsychINFO database to perform the search and the search terms became slightly modified in the process.

<sup>10</sup>The search of C2 Spectr was conducted in February 2006 at the recommendation of Campbell Collaboration reviewers of this product.

<sup>11</sup>Zief and Maynard coded this study in order to avoid a conflict of interest and determined that this study met the inclusion criteria for this review.

Table 3: Experimental Design Studies Identified Through Prior Reviews

Author	Program	Inclusion decision
Hahn, Leavitt, and Aaron (1994)	Quantum Opportunities	No—Significant summer and comprehensive services component
LoSciuto, Freeman, Harrington, Altman, and Lamphear (1997)	Woodrock Youth Development	No—Significant school-day component
Phillips, R. (1999)	No name	No—Special needs population
Weisman, Soule, Womer, and Gottfredson (2001)	Maryland After-School Community Grant Program 4	Yes—academic, recreational, and youth development programming
Weisman et al. (2001)	Maryland After-School Community Grant Program 17	Yes—academic, recreational, and youth development programming
Weisman et al. (2001)	Maryland After-School Community Grant Program 21	Yes—academic, recreational, and youth development programming

**Internet searches.** The internet search generated almost three million citations (Table 4); the researchers initially reviewed 400 of these citations, and all were either irrelevant or duplicative of studies identified through other sources. Therefore, the researchers did not feel that devoting additional time and resources to the remaining citations would have yielded any additional studies that would have met the inclusion criteria for this review.

Table 4: Internet Searches and Results

Google.com search terms <sup>12</sup>	Date	Citations	Other parameters
<i>With all the words</i> + after school program evaluation	September 9, 2003	2,380,000	English language Any format Anytime
<i>With all the words</i> + after school program experimental evaluation	September 9, 2003	556,000	Anywhere on page No filtering

### Screening of Studies for Inclusion in the Review

Both primary reviewers reviewed the abstracts of all studies appearing eligible and applied a first screen to identify relevant studies (i.e., type of after-school program, characteristics of target population, type of study, and outcome measures). Studies that appeared to meet these basic criteria were retrieved and further assessed. Those studies that met the established inclusion criteria were held for an assessment of study quality. If there had been disagreement about whether to retain a study for further assessment at this point, the advisor to the reviewers, Dr. Rebecca Maynard, would have been consulted. However, there were no disagreements about whether a study met the preliminary criteria for inclusion.

<sup>12</sup>Google.com recommended inserting “+” before the search string because “after” could be identified as a stop term by the search engine.

***Judging study quality.*** The reviewers developed a formal approach to assess the quality of the relevant experimental design studies identified (Appendix 3). From the list of the 39 qualities examined that the reviewers felt should be included in the reporting of a rigorous experimental study, four specific standards needed to be met in order for a study to be included in this review: (a) no specific evidence of control group contamination, (b) neither overall study attrition nor differential attrition would bias the impact estimates, (c) appropriate statistical measures were used for the analyses, and (d) the primary impact analyses were conducted on all available sample members at followup (i.e., “intention to treat” and not “treatment on treated”).

Both reviewers independently read each full study, applied the study quality criteria, and recommended whether the study should be included in the final review. In many cases, authors were contacted and asked to provide information that was considered critical to assess a study’s quality; all contacted authors provided such information, and no studies were rejected due to poor study quality. If the reviewers had disagreed on their assessment of a study’s quality, the advisor to the reviewers would have been consulted. However, this was not necessary.

### Extracting Data from Included Studies

Both reviewers captured key data from each of the five studies included in the review using an internally designed coding form and then entered this data in an Access database.<sup>13</sup> The data entered for each item were compared and any coding differences were resolved by both authors independently returning to the study in question. No coding differences remained after this process.

### Analytic Methods

The impact estimates presented by the five studies were analyzed in two primary ways. First, all measured outcomes in each outcome area (context, activities, behavior, social and emotional, academic) were tabulated, including indications of the direction and statistical significance of program impacts. This descriptive analysis is not meant to suggest overall program impacts as each single data point is heavily dependent on the sample size of the study from which it originated. Instead, it provides a basis for discussing what was measured and in which areas significant and null outcomes are found.

Next, for all similarly measured outcomes, meta-analytic techniques were applied using Comprehensive Meta-Analysis software by Biostat.<sup>14</sup> When the meta-analysis involved the difference in means (i.e., test scores, hours in self-care, hours watching television, and GPA) a

---

<sup>13</sup>Please contact the review authors for a copy of the coding form.

<sup>14</sup>The Comprehensive Meta-Analysis software by Biostat used for the meta-analytic calculations in this review requires standard deviations when calculating the standardized mean difference (Cohen’s *d*). In cases where the standard deviations were not reported by the study authors, we were able to compute a pooled standard deviation using the following method:

1. Using the reported *p*-value and total sample size for the specific outcome, we identified the *t*-statistic.
2. Using Formula 15, Table B10 in Lipsey and Wilson (2001, p. 200), we calculated the pooled standard deviation for the outcome.
3. The pooled standard deviation was then used to calculate the effect size.

standardized mean difference (Cohen's *d*) was calculated by the software. For those outcomes reported as the proportion of youth experiencing an event (i.e., participating in art/music/drama/dance or athletic activities), a standardized rate difference was calculated by the software.<sup>15</sup> In order to ensure independence, each study only contributed one effect to each calculation. A fixed-effects model was used for all meta-analytic calculations because this model allows for inference only to this small number of identified studies. In this way, we are addressing the question of how *these* studied programs impact student outcomes and are not making inferences to a larger group of studies not yet identified or conducted.

For each meta-analysis conducted, a Forrest Plot displays the differences between each study's individual effect, the overall effect for the outcome, and the confidence intervals around the point estimates. In general, the narrower the confidence interval, the more confident we are that the true impact is near the estimated value. The width of the confidence interval is inversely related to the sample size.<sup>16</sup>

It is commonly held that studies published in peer-reviewed journals are easier than unpublished work to identify and retrieve, and therefore a review without an extensive search strategy might have a disproportionate representation of published reports. This is a concern because published work may have a disproportionate number of studies showing statistically positive effects. It is important to consider such bias upon analysis and this is often done by calculating the *fail-safe N*, a comparison of the mean effects from published and unpublished studies. None of the included studies in this review were published in peer-reviewed journals, and therefore we did not investigate potential publication bias in this way.<sup>17</sup>

### Characteristics of Evaluated Programs

Along many dimensions, the five evaluated programs included in this review are similar (Table 5).<sup>18</sup> Most operated in urban, school-based environments and served primarily low-income minority students in poor-performing schools. The principal program funding source was the federal government. Three of the five studies were of programs designed primarily to reduce negative behaviors (e.g., delinquency and drug use), and all five programs included activities to promote positive youth outcomes, like academic growth and exposure to enriching activities in a

---

<sup>15</sup>While this method does provide an estimate of the mean proportion across studies, it can underestimate the size of the confidence interval around the mean effect size (proportion) and overestimate the degree of heterogeneity across the effect sizes (Lipsey & Wilson, 2001). This is especially true when the observed proportions are less than .2 or greater than .8. However, the studies included in this review did not present proportions at these two extremes. Additionally, this analysis was determined to be sufficient for this exploratory review of the available experimental research.

<sup>16</sup>A funnel plot is another useful graphic in meta-analyses to show effect size by sample size, which can show potential bias due to under representation of studies with small samples. However, due to the small number of studies included in this review, a funnel plot would have not aided the analysis.

<sup>17</sup>A *fail-safe N* is also referred to as a *Rosenthal fail-safe*.

<sup>18</sup>The elementary experimental component of the national study of 21<sup>st</sup> CCLCs had an experimental design whereas the middle school component had a quasi-experimental design. For the purpose of this review, only the experimental impact estimates from the elementary component are reported; the included data represent seven elementary centers, and because the data were not provided separately for each center, the elementary component of the 21<sup>st</sup> CCLC elementary evaluation is included as a single study in this review (U.S. Department of Education, 2003).

safe environment. The programs all used a similar mix of activities, combining academic support services with recreational and, in some cases, youth development programming. Program youth primarily had exposure to the intervention for an entire school year, but one program was in operation for only five months. A more detailed description of the program and evaluation features for the included studies is found in Appendix 4.

The largest of the evaluations was conducted by a nationally prominent research firm; the other, smaller scale studies were conducted by university research teams. The studies used similar data sources—surveys and school records—to measure a wide variety of outcomes. However, there was little convergence among specific outcome measures among the five studies. In all cases, follow-up data collection occurred at the end of the program year. Not one study provided a detailed description of the activities of the control group, yet there was no evidence that the control group received the intervention or a similar treatment.

Three of the evaluated programs were of the Maryland After-School Community Grant Program (MASCGP), were conducted by the same research team, received funding from the same source, and had the same program goals (Weisman et al., 2001). Despite these common features, there were some differences in the program models and implementation, and researchers disaggregated the data by site. For these reasons, each program’s evaluation is considered as a separate study in this review.

## **Impacts on Student Outcomes**

The studies included in this review measured outcomes that cluster in five key areas—student context (i.e., student location, supervision, and safety); participation in enriching activities; and behavioral, social and emotional, and academic outcomes. A total of 97 different outcomes were measured, and 79 percent of these outcomes were measured by only one study. Even when more than one study measured the same outcome, the measurements mostly came from different, internally designed surveys or were reported by different groups (students, parents, or teachers).

The findings from these studies are grouped in the following discussion by the five primary outcome areas and presented according to a theoretical trajectory of student changes as a result of participation in an after-school program (see Appendix 1)—changes in context and increased participation in enriching, engaging activities are precursors of improved behavioral, social and emotional, and academic outcomes. The discussion of findings for each outcome area is first based on “vote counting” tables which show the items measured for each outcome area, the distribution of the significant findings, and the number of null impacts.<sup>19</sup> Overall, a great majority of the measured outcomes revealed null impacts. Finally, any meta-analyses that could be conducted are presented.

---

<sup>19</sup>For ease of interpreting the 97 different outcomes, we created sub-groups under each outcome area which included related outcome measures. For example, under *behavioral* outcomes, absenteeism and lateness have been combined as measures of *school attendance*; helping another student in school, helping another student after school, getting along with others, and working in a group combine as outcomes showing students’ *cooperation*.

Table 5: Characteristics of the Evaluated Programs Included in this Review

Feature	Total (%)
Setting	
Urban	60%
Rural	0%
Suburban	20%
Multiple	20%
Principal Program Funding Source	
Federal government	80%
Local school district	20%
Intervention Location	
Community organization	20%
School	80%
Grade Level	
Elementary school	80%
Middle school	20%
High school	0%
Service Delivery	
Community organization	20%
School	80%
Primary Program Goal	
Reduce negative behaviors	60%
Promote positive outcomes	100%
Intervention Components	
Academic	100%
Recreation	100%
Developmental	60%
Duration of Intervention	
5 months	20%
9 months	80%
Period of Followup	
Immediate end of intervention	100%
Sample Size	
Less than 100	60%
Between 100 and 250	20%
Greater than 250	20%

The incongruent measures between the studies limited the possibilities for meta-analysis. The meta-analyses that were able to be conducted (with sometimes as few as two studies) should not be used to identify whether after-school programs are or are not meeting their goals. Although the studies present credible estimates of program impacts and the programs are fairly homogeneous in terms of goals and activities, they represent a small proportion of the programs in operation and have collected data over a short time period. Still, it is of interest to note where the greater and more significant effects were found and what they may suggest for future research and areas to consider for program improvement (Table 6).

Table 6: Results of Meta-Analysis

Outcome Measure	Studies	N	Program duration (months) <sup>a</sup>	Effect size estimate	
				Rate difference	Standardized mean difference (Cohen's <i>d</i> )
<b>Context</b>					
Self-care	Weisman et al. (3 program evaluations)	101	9		.503 **
<b>Activities</b>					
Athletic activities	Lauver U.S. Department of Education	661	5 9	.073 *	
Art/music/drama dance activities	Lauver U.S. Department of Education	661	5 9	.083 *	
<b>Behavior</b>					
School attendance	Lauver U.S. Department of Education	1,072	5 9		-.001
Television viewing	Lauver U.S. Department of Education	661	5 9		-.066
<b>Academics</b>					
Reading scores	Lauver U.S. Department of Education	983	5 9		.028
Grade Point Average (GPA)	Lauver U.S. Department of Education Weisman et al. (3 program evaluations)	1,004	5 9 9		.083

<sup>a</sup> In all cases, the administration of the post-test occurred immediately at the end of the intervention year.

\*\* =  $p \leq .01$ , \* =  $p \leq .05$

### Student Context

The recent interest in after-school programs reflects the growing concern over the number of children in self-care daily and the negative effects of being unsupervised between the hours of 3:00 – 6:00 p.m. The mechanisms through which after-school programs aim to improve outcomes for participants is by changing the environments in which youth spend their after-school time—for example, increasing time in safe, supervised settings; creating more positive peer associations; and increasing parental involvement in home and school activities. Across the five studies in this review, 21 contextual factors were measured (Table 7).

The three experimental studies of MASCGP programs measured program-related differences in the quality of peer associations. There is no evidence of favorable program impacts on peer associations among the three programs for which this outcome was measured. In fact, participants in MASCGP Program 21 were less likely to be in the company of positive peers ( $p = .01$ ) (Weisman et al., 2001). The researchers suggest that such negative impacts could be the result of operational decisions to provide a minimum level of social skills activities and to allow program participants to self-monitor their social skills during recreation activities.

Table 7: Studies' Reported Program Impacts on Student Context

Outcome area	# Studies	# Measured outcomes	Outcomes favoring treatment		Outcomes favoring control		Null impacts	
			$p \leq .05$	$.05 < p \leq .10$	$p \leq .05$	$.05 < p \leq .10$	#	Percentage
Peer associations <sup>a,e</sup>	3	6	0	0	1	0	5	83%
Parent involvement <sup>a,b</sup>	1	7	2	1	0	0	4	57%
Safety <sup>a,c</sup>	1	3	0	0	0	0	3	100%
Student location after school <sup>a,d</sup>	1	5	3	0	0	0	2	40%
<b>Supervision<sup>a,f</sup></b>	<b>5</b>	<b>11</b>	<b>3</b>	<b>0</b>	<b>0</b>	<b>3</b>	<b>5</b>	<b>45%</b>
Total	5	32	8	1	1	3	19	59%

Sources: Lauver (2002); U.S. Department of Education (2003); Weisman et al. (2001).

Note: Category in bold is included in meta-analysis.

<sup>a</sup> Outcome measured multiple ways within a single study

<sup>b</sup> Includes helped child with homework, checked homework, asked child about class, attended an after-school event, attended a PTO meeting, attended a school open house, and volunteered at school

<sup>c</sup> Includes student feels not at all safe, student feels somewhat safe, and student feels very safe

<sup>d</sup> Includes mean days after school for activities and percentage in the following three or more days per week: (not in) own home, someone else's home, at school or other place for activities, and mixed location for activities

<sup>e</sup> Includes time with positive peers, decreased time with negative peers

<sup>f</sup> Includes non-parent supervision, parent supervision, days in sibling care, days per week in self-care, daily hours in self-care, and hours per week in self-care

Only the 21<sup>st</sup> CCLC study measured program-related differences in parent involvement, child safety, and location where students spend their after-school time (U.S. Department of Education, 2003). Specifically, the elementary component of the national evaluation of 21<sup>st</sup> CCLCs found the following differences:

- (a) Parents of participants were more likely to help their children with homework ( $p = .02$ ), ask their child about class ( $p = .07$ ), and attend an after-school event ( $p = .04$ ).
- (b) Participants stayed after school for more days in a typical week ( $p = 0$ ), were at school or another place for activities three or more days in the previous week ( $p = .02$ ), were more likely to be in the care of an adult who was not a parent ( $p = .01$ ), were less likely to be in their own home ( $p = .01$ ), were less likely to be under parental supervision during the after-school hours ( $p = .09$ ), and were also less likely to be in sibling care ( $p = .09$ ).

Researchers of this study concluded that while the program changed student location and who was providing supervision, it did not reduce incidences of self-care (U.S. Department of Education, 2003). In sum, of the 11 supervision impact estimates reported from all five studies included in this review, five were not statistically significant, three favored the treatment group, and three favored the control group.

**Meta-analysis on hours per week in self-care.** Given the non-comparable measurements of student supervision, it was only appropriate to combine the impact estimates from the three MASCGP sites. The pooled results indicate an overall moderate, positive effect ( $d = .503$ ,  $p = .03$ ; Tables 8 and 9) on the number of hours a student spent in self-care. The distribution of the effect sizes contributed from each study is fairly homogeneous, signaling that the dispersion is no greater than would be expected from sampling error alone ( $Q = 1.12$ ,  $df = 2$ ,  $p = .57$ ).

Overall, the control youth spent an average of 4.0 hours per week in self-care while the treatment youth spent an average of 1.53 hours per week in self-care.<sup>20</sup> These three small, similar programs appear to be reducing self-care by 2.5 hours per week. However, it would not be appropriate to generalize beyond these findings and conclude that after-school programs are effective at reducing hours spent in self-care.

Table 8: Hours per Week in Self-Care: Standardized Mean Difference ( $d$ )

Citation	Effect	Lower	Upper	-2.00	-1.00	0.00	1.00	2.00
MASCGP Study 17	.748	-.136	1.632					
MASCGP Study 21	.241	-.452	.934					
MASCGP Study 4	.598	-.052	1.249					
<b>Fixed Combined (3)</b>	<b>.503</b>	<b>.097</b>	<b>.910</b>					

<sup>20</sup>This calculation involves multiplying the effect size by the weighted, average control standard deviation and then adding that value to the weighted, average hours unsupervised for the control youth. The degrees of freedom (n-1) are used for the weights. A more thorough discussion of this can be found in Lipsey and Wilson (2001, pp. 149-150).

Table 9: Mean Hours per Week in Self-Care Reported by Studies

Study	<i>N</i>	Program youth	Control youth	Difference	<i>p</i> -value
MASCGP 17	24	1.73	7.35	5.62	.02
MASCGP 21	35	2.41	3.39	0.98	.47
MASCGP 4	42	0.89	2.78	1.89	.03

### Participation in Activities

Students in after-school programs that offer a variety of enriching activities, such as the programs included in this review, may have more of an opportunity to participate in enriching activities than their peers not attending after-school programs. Such activities are potentially important for further engaging youth in after-school programming, diminishing opportunities for risk-taking behaviors, and recognizing and developing youth skills and talents. Furthermore, engaging youth in after-school programs through the use of recreational activities provides opportunities for improving behaviors and academic performance.

Only two of the studies included in this review—Lauver (2002) and U.S. Department of Education (2003)—examined the extent to which there were measured differences between the program and control group’s participation in activities; they reported 16 impact measures (Table 10). Lauver (2002) found that participants were more likely than their control group counterparts to participate in strength training ( $p = .001$ ). This finding reflects the primary goal of the program and a specific program component. Elementary school youth who participated in the experimental study of 21<sup>st</sup> CCLCs were more likely than their control group counterparts to report caring for a younger sibling ( $p = .01$ )<sup>21</sup> and less likely to participate in clubs like Boys and Girls Clubs or Boy and Girl Scouts ( $p = .05$ ; U.S. Department of Education, 2003). The lower club participation suggests that after-school programs may supplant, and not supplement, other opportunities for low-income youth.

***Meta-analysis on participation in athletic and art/music/drama/dance activities.*** The results of the completed meta-analyses show a very small but significant effect on students’ participation in athletic activities (*rate difference* = .073,  $p = .04$ ; Tables 11 and 12).<sup>22</sup> However, there should be caution in interpreting these results as there is considerable heterogeneity in the estimated impacts across these two studies ( $Q = 5.39$ ,  $df = 1$ ,  $p = .002$ ). Data from the same two studies suggests that participation in the programs is having a very small, significant effect on youth participation in art/music/drama/dance activities (*rate difference* = .083,  $p = .02$ ; Tables 13 and

<sup>21</sup>Although we have labeled this as an impact favoring treatment youth, it is not entirely clear how to interpret the higher rate of sibling care. It may signal that, if siblings of experimental youth were also given access to the program, the older sibling would have care-taking responsibilities for the younger sibling on the way to or home from the program.

<sup>22</sup>Lauver (2002) measured participation in athletic activities two ways—percent of students who spent one or more hours a week strength training and percent who spent one or more hours per week in exercise or sports. It is likely that a number of students would have been counted twice if both of these outcomes were used to calculate the effect of participation in athletic activities from this one study. Therefore, the measure that showed the highest impact contributed to the effect size estimate—percent who spent one or more hours a week strength training. This approach will provide an upper-bound estimate of the pooled effects. The national evaluation of 21<sup>st</sup> CCLC reported one outcome in this area—percent of students who participated in organized sports the prior day after school.

Table 10: Studies' Reported Program Impacts on Participation in Enriching Activities

Outcome Area	# Studies	# Measured outcomes	Outcomes favoring treatment		Outcomes favoring control		Null impacts	
			$p \leq .05$	$.05 < p \leq .10$	$p \leq .05$	$.05 < p \leq .10$	#	Percentage
<b>Enrichment activities</b> <sup>a,b</sup>	<b>2</b>	<b>4</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>3</b>	<b>75%</b>
Hanging out	1	1	0	0	0	0	1	100%
Home/family responsibilities <sup>a,c</sup>	1	2	1	0	0	0	1	50%
Non-homework academic activities <sup>a,d</sup>	1	4	0	0	0	0	4	100%
<b>Physical activities</b> <sup>a,e</sup>	<b>2</b>	<b>4</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>3</b>	<b>75%</b>
Religious activities	1	1	0	0	0	0	1	100%
Total	2	16	2	0	1	0	13	81%

Sources: Lauver (2002); U.S. Department of Education (2003).

Note: Categories in bold are included in meta-analysis.

<sup>a</sup> Outcome measured multiple ways within a single study

<sup>b</sup> Includes art, music, dance activities; art, music, dance lessons; clubs; and school activities

<sup>c</sup> Includes care of sibling, chores around house

<sup>d</sup> Includes computer activities; tutoring; reading for fun; and reading, writing, and science activities

<sup>e</sup> Includes play, strength training, sports, and time exercising

14).<sup>23</sup> Yet here the pooled results show a more homogeneous distribution, signaling that the dispersion is no greater than would be expected from sampling error alone ( $Q = .00004$ ,  $df = 1$ ,  $p = .98$ ).

Table 11: Athletic Activities: Rate Difference

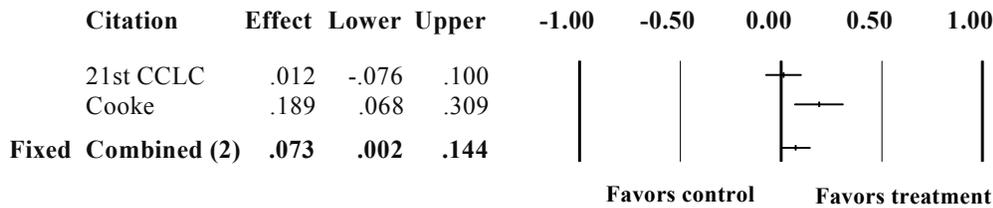


Table 12: Proportion of Youth in Athletic Activities Reported by Studies

Study	<i>N</i>	Program youth	Control youth	Difference	<i>p</i> -value
21 <sup>st</sup> CCLC	441	28.7	27.3	1.4	.80
Cooke	220	79.7	61.1	18.6	.00

Table 13: Art/Music/Drama/Dance Activities: Rate Difference

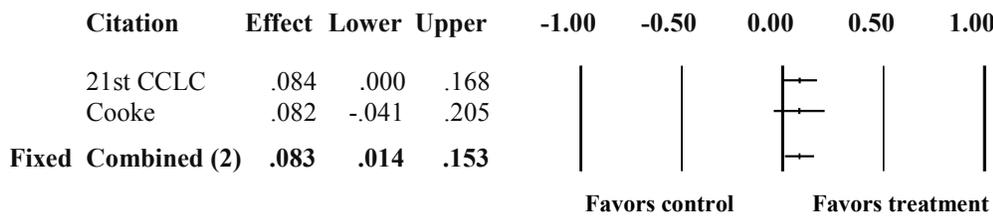


Table 14: Proportion of Youth in Art/Music/Drama/Dance Activities Reported by Studies

Study	<i>N</i>	Program youth	Control youth	Difference	<i>p</i> -value
21 <sup>st</sup> CCLC	441	30.2	21.9	8.4	.11
Cooke	220	72.8	64.8	8.0	.24

## Behavior

The five studies measured a total of 36 behavioral outcomes (Table 15). One positive significant outcome for program participants is found: a greater percentage of participants in the Cooke Middle School After-School Recreation program spent one or more hours on homework each

<sup>23</sup>The national evaluation of 21<sup>st</sup> CCLC measured art/music/drama/dance outcomes two ways: the percent of students in school activities (band, music, drama, etc.) and percent of students in lessons (music, art, dance) (U.S. Department of Education, 2003). The measure that showed the highest impact was used for the meta-analysis—percentage of students in lessons (music, art, dance, etc.). This approach will provide an upper-bound estimate of the pooled effects. Lauver (2002) reported one outcome in this area—percent of students who participated in one or more hours of art, music, or dance per week.

Table 15: Studies' Reported Program Impacts on Behavioral Outcomes

Outcome area	# Studies	# Measured outcomes	Outcomes favoring treatment		Outcomes favoring control		Null impacts	
			$p \leq .05$	$.05 < p \leq .10$	$p \leq .05$	$.05 < p \leq .10$	#	Percentage
<b>Attendance</b> <sup>a,b</sup>	<b>2</b>	<b>3</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>3</b>	<b>100%</b>
Attitudes towards risk-taking behaviors	3	3	0	0	0	0	3	100%
ATD use	2	2	0	0	0	0	2	100%
Overall behavior	2	2	0	0	0	0	2	100%
Cooperation <sup>a,c</sup>	1	4	0	1	0	0	3	75%
Delinquency	3	3	0	0	0	0	3	100%
Disciplinary actions <sup>a,d</sup>	1	5	0	0	0	0	5	100%
Effort <sup>a,c</sup>	1	3	0	1	0	1	1	33%
Homework <sup>a,f</sup>	2	4	1	0	0	0	3	75%
Ready to learn	1	1	0	0	0	0	1	100%
Social skills	3	3	0	0	0	0	3	100%
<b>Television watching</b> <sup>a,g</sup>	<b>2</b>	<b>3</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>3</b>	<b>100%</b>
Total	5	36	1	2	0	1	32	89%

Sources: Lauver (2002); U.S. Department of Education (2003); Weisman et al. (2001).

Note: Categories in bold are included in meta-analysis.

<sup>a</sup> Outcome measured multiple ways within a single study

<sup>b</sup> Includes absenteeism, lateness

<sup>c</sup> Includes helping another student in school, helping another student after school, getting along with others, working with others in a group

<sup>d</sup> Includes call to parent, detention, discipline for misbehaving, suspension, and being sent to office for misbehaving

<sup>e</sup> Includes overall effort composite score, tries hard in English/reading, works hard at school

<sup>f</sup> Includes completion frequency, satisfactory completion, time on homework, homework yesterday (Y or N)

<sup>g</sup> Includes time watching TV, watched TV yesterday

day ( $p = .04$ ; Lauver, 2002). The 21<sup>st</sup> CCLC evaluation reported two behavioral outcomes that may be approaching significance: elementary participants reported that they were more likely to help another student after school ( $p = .07$ ), and their teachers reported that they were more likely to try hard in English/Reading ( $p = .08$ ; U.S. Department of Education, 2003). The same study found a possible negative, behavioral effect—the parents of treatment youth were less likely to agree that their child worked hard at school ( $p = .06$ ; U.S. Department of Education, 2003).

**Meta-analysis on school attendance and television viewing.** Only school attendance and television viewing were measured by more than one study in a way that permitted pooling of individual effects. These behavioral outcomes are of particular interest to policymakers and the public as an understandable link can be made between improving these behaviors and student academic success. If students become interested in and engaged with their after-school program, it reasonably follows that their school attendance may increase as they go to school more regularly in order to participate in the program at the end of the day. However, the pooled effects show no measured impact on school attendance ( $d = -.001$ ,  $p = .98$ ; Tables 16 and 17). The distribution of the effect sizes contributed from each study is fairly homogeneous, signaling that the dispersion is no greater than would be expected from sampling error alone ( $Q = .345$ ,  $df = 1$ ,  $p = .55$ ).

Table 16: School Attendance: Standardized Mean Difference ( $d$ )

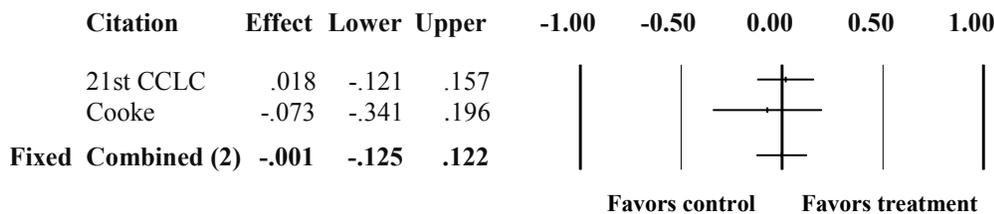


Table 17: Mean Days Absent Reported by Studies

Study	$N$	Program youth	Control youth	Difference	$p$ -value
21 <sup>st</sup> CCLC	854	8.2	8.2	0.0	.91
Cooke	218	24.0	22.5	-1.5	.63

The amount of time children spend watching television is a major concern of policymakers and the public. Many argue that television exposes children to violence and other inappropriate activities for youth and decreases time from homework and other instructional activities like reading. In theory, participation in after-school programs could reduce the number of hours a student spends watching television. The estimated overall impact on television viewing from two studies in this review is not positive, nor is it statistically significant ( $d = -.066$ ,  $p = .41$ ; Tables 18 and 19). The distribution of the effect sizes contributed from each study is fairly homogeneous, signaling that the dispersion is no greater than would be expected from sampling error alone ( $Q = .052$ ,  $df = 1$ ,  $p = .82$ ).

Table 18: Television Viewing: Standardized Mean Difference (*d*)

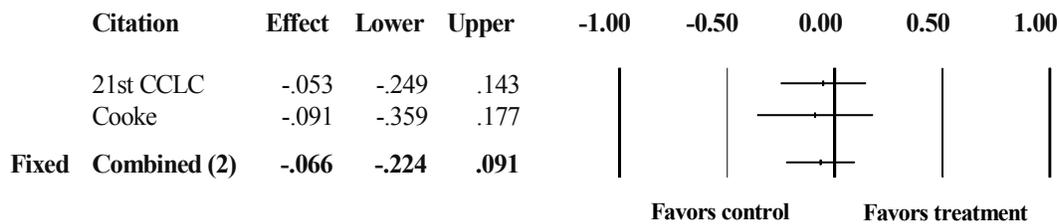


Table 19: Mean Hours of Television Viewing Reported by Studies

Study	<i>N</i>	Program youth	Control youth	Difference	<i>p</i> -value
21 <sup>st</sup> CCLC	220	4.3	4.0	-0.3	.51
Cooke	441	2.0	1.9	-0.1	.63

### Social and Emotional

In theory, participation in a supportive environment with caring adults and peers could promote positive social/emotional outcomes for youth, such as improved bonds with adults and peers and greater future aspirations. The five studies measured a total of 11 social and emotional outcomes (Table 20). Only one significant social and emotional outcome was found: participants had greater college aspirations than non-participants ( $p = .01$ ; Lauver, 2002). The complementary process evaluation of this program suggests that the bond that developed between a young college volunteer at the program and a group of participants may have influenced this impact finding and may speak for the need to staff programs with older students who share similar background characteristics with participants and can be role models for the younger youth.

The social and emotional outcomes measured by the five studies that showed null effects included feeling left out, sticking to beliefs, feeling bad for others, believing the best about people, and bonding. No outcomes were measured in a way that allowed for pooling of effects across studies.

### Academics

After-school programs are increasingly recognized by policymakers as a possible means to boost participants' academic outcomes, especially for those students considered more academically at-risk—low-income minority youth in poor-performing urban schools. The increase in federally funded after-school programs through 21<sup>st</sup> CCLC appropriations has increased the supply of programs offering academic programming. Furthermore, the demand from parents for programs to include academic support might reflect the decreased time working parents have to help with homework and remediation.

*Of the 13 estimates of program impacts on academic outcomes measured by the five studies, one significant effect was found (Table 21): elementary participants had significantly higher social studies grades than control youth (U.S. Department of Education, 2003). It is important to*

Table 20: Studies' Reported Program Impacts on Social and Emotional Outcomes

Outcome area	# Studies	# Measured outcomes	Outcomes favoring treatment		Outcomes favoring control		Null impacts	
			$p \leq .05$	$.05 < p \leq .10$	$p \leq .05$	$.05 < p \leq .10$	#	Percentage
Aspirations	1	1	1	0	0	0	0	0%
Believing the best about people	1	1	0	0	0	0	1	100%
Bonding <sup>a,b</sup>	3	6	0	0	0	0	6	100%
Feeling bad for others	1	1	0	0	0	0	1	100%
Feeling left out	1	1	0	0	0	0	1	100%
Sticking to beliefs	1	1	0	0	0	0	1	100%
<b>Total</b>	<b>5</b>	<b>11</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>10</b>	<b>91%</b>

Sources: Lauver (2002); U.S. Department of Education (2003); Weisman et al. (2001).

<sup>a</sup> Outcome measured multiple ways within a single study

<sup>b</sup> Includes involvement, attachment, and commitment

Table 21: Studies' Reported Program Impacts on Academic Outcomes

Outcome area	# Studies	# Measured outcomes	Outcomes favoring treatment		Outcomes favoring control		Null impacts	
			$p \leq .05$	$.05 < p \leq .10$	$p \leq .05$	$.05 < p \leq .10$	#	Percentage
<b>Grades<sup>a,b</sup></b>	<b>5</b>	<b>8</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>7</b>	<b>88%</b>
Knowledge of computer use	1	1	0	0	0	0	1	100%
Reading confidence	1	1	0	0	0	0	1	100%
Standardized math test scores	1	1	0	0	0	0	1	100%
<b>Standardized reading test scores</b>	<b>2</b>	<b>2</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>2</b>	<b>100%</b>
<b>Total</b>	<b>5</b>	<b>13</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>12</b>	<b>92%</b>

Sources: Lauver (2002); U.S. Department of Education (2003); Weisman et al. (2001).

Note: Categories in bold are included in meta-analysis.

<sup>a</sup> Outcome measured multiple ways within a single study

<sup>b</sup> GPA based on grades in math, English/language arts, science, and social studies

consider program goals when interpreting academic outcomes such as test scores and grade point averages. All studies included in the review contained an academic component, ranging from homework help to programs in which improving students' academic performance was identified as a major goal of the program.

**Meta-analysis on reading achievement and grades.** Pooling the effect sizes from the two studies reporting standardized reading scores suggests that these two programs have not had an effect on reading achievement ( $d = .028, p = .67$ ; Tables 22 and 23). The distribution of the effect sizes contributed from each study is fairly homogeneous, signaling that the dispersion is no greater than would be expected from sampling error alone ( $Q = .16, df = 1, p = .69$ ).

Table 22: SAT-9 Reading Percentile: Standardized Mean Difference ( $d$ )

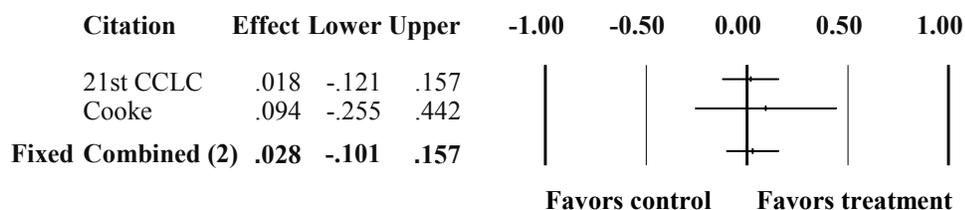


Table 23: Mean SAT-9 Reading Percentiles Reported by Studies

Study	$N$	Program youth	Control youth	Difference	$p$ -value
21 <sup>st</sup> CCLC	854	34.3	34.2	0.1	.96
Cooke	129	33.9	31.7	2.2	.62

Pooling the measured effects of program participation on grade point averages reveals a different story than viewing each of the study's measured impacts separately. The meta-analysis reveals that programs are having a small but insignificant effect on improving participants' grades ( $d = .083, p = .16$ ; Tables 24 and 25).<sup>24</sup> The distribution of the effect sizes contributed from each study is fairly homogeneous, signaling that the dispersion is no greater than would be expected from sampling error alone ( $Q = 4.37, df = 4, p = .36$ ).

<sup>24</sup>In order to prepare the data for this analysis, the reported grade point averages from all five studies were converted to a 100-point scale. The 21<sup>st</sup> CCLC study did not report an overall GPA but instead separately reported the GPA from four subjects—math, English/language arts, science, social studies/history (U.S. Department of Education, 2003). An overall GPA from this study was calculated by averaging the four individual subject GPAs. The four standard deviations were also averaged.

Table 24: Grade Point Average: Standardized Mean Difference (*d*)

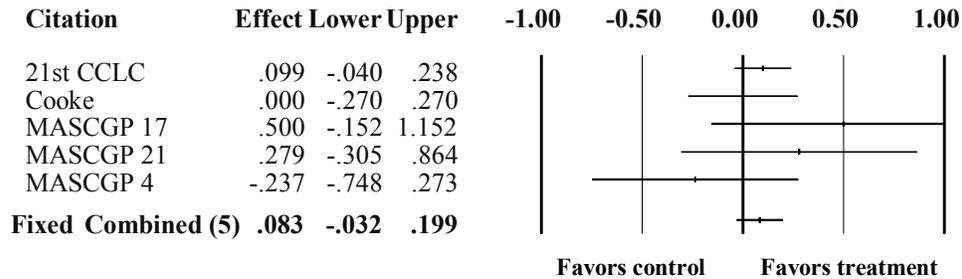


Table 25: Mean Grade Point Average Reported by Studies

Study	N	Program		Difference	<i>p</i> -value
		youth	Control youth		
21 <sup>st</sup> CCLC	854	82.7	81.3	1.4	.16
Cooke <sup>a</sup>	216	80.0	80.0	0.0	.64
MASCGP 17	40	75.6	73.2	2.4	.12
MASCGP 21	48	84.4	82.3	2.1	.34
MASCGP 4	62	80.6	82.4	-1.8	.34

Note: <sup>a</sup> Cooke GPA is from fourth marking period (Lauver, 2002).

## Findings

This review sought to answer the following questions with the available high quality experimental studies of programs that combined academic programming with other activities such as youth development or recreation:

1. To what extent and in what ways does access to after-school programs impact student context (i.e., student location, supervision, and safety), participation in enriching activities, behaviors, social and emotional development, and academic outcomes for youth?
2. Do the effects of after-school programs vary among subgroups of youth defined by their baseline characteristics?
3. Among the program models and settings evaluated, do some seem more beneficial to youth than others? What are the distinguishing characteristics of those more and less successful programs?

An extensive search of the literature uncovered only five studies that met the inclusion criteria for this review. The five evaluated programs were fairly homogenous along many dimensions, including their target populations and settings—primarily elementary youth living in lower-income, urban settings. These similarities prevented an analysis of the variation in impacts by youth characteristics or program settings.

The programs themselves placed different emphases on promoting positive behaviors and reducing negative ones. However, it does not appear that any one approach was more or less effective at contributing towards improved behavioral outcomes or other estimated effects. The studies included in this review also provide no evidence that any one program model is more effective at changing students' context or improving academic outcomes.

A complementary process evaluation of the Cooke Middle School After-School Recreation Program revealed some aspects of that program model were successful at improving future aspirations and participation in activities (Lauver, 2002). Program youth had significantly higher college aspirations than nonparticipants, and this finding suggested that the relationship participants developed with a college student volunteer at the program who shared many background characteristics with the youth was important. Furthermore, the program had a significant, positive impact on increasing youth participation in strength training activities, a primary goal of the program and consistent program feature.

The estimated pooled effects from two program evaluations included in this review suggest that participation positively influenced the participation in athletic and art/music/dance/drama activities. While the effects are small (less than .10 of a standard deviation), they are significant ( $p = .04, .02$ , respectively). This finding is encouraging as such changes are possibly necessary to support participants' further growth and development.

The pooled results from three after-school programs revealed that the incidence of self-care decreased by more than .5 standard deviations ( $p = .03$ ). It is important to consider that these findings are pooled from three small sites of one program—the Maryland After-School Community Grant Program—and should not be generalized to other programs. An additional data from one of the studies revealed that the program had a significant, negative effect on peer associations. Programs may need to pay attention to the kind of supervision they are offering—students may be supervised but are not in an environment conducive to promoting positive behaviors.

While the overall average standardized reading test scores show that the programs did not contribute to higher reading achievement for participants ( $d = .028, p = .67$ ), after-school programs may be having more of an impact on raising grades. However, these impacts were quite small and not significant ( $d = .083, p = .16$ ). Yet they are of interest given that they present the second largest measured pooled effects. A logic model for understanding the mechanisms for changing outcomes as a result of participation in after-school programs suggests that higher grades might occur after changes in students' behaviors and social and emotional outcomes (see Appendix 1). However, our analysis shows stronger effects for improved grades than for the behaviors that could impact grades, like improved school attendance and decreased television viewing.

Notably, looking across the 97 impacts measured by the five studies included in this review reveals primarily null findings—84 percent showed no significant differences between the program and control youth. Also, not one of these studies reported impacts for parents. In this era of welfare reform, it might be important to understand how this type of intervention targeting low-income youth may impact parents' job attendance and retention or parental levels of stress

while balancing the demands of work and child care. Such parental outcomes could arguably mediate student social and emotional, behavioral, and academic outcomes.

### Comparison of Findings Between this Review and Prior Reviews

The inclusion criteria for this review differ significantly from other prior reviews in the field, particularly in the more specific definition of the intervention and the inclusion of only well-implemented experimental design studies (Table 26). Additionally, prior reviews have relied solely on a narrative description of the findings and/or listing of only significant impacts and have not employed the more robust meta-analytic techniques to understand overall program effects.

Similar to this review, those reviews that included the national evaluation of 21<sup>st</sup> CCLC also suggested that programs may contribute to increased grades (Kane, 2004; Little & Harris, 2003). However, their findings were based on the significant program impact on improving social studies grades from this one study, not a meta-analysis of overall GPA using data from five evaluations (the more robust technique used in this review). Alternatively, the findings from one review suggested that programs contributed to improved achievement test scores and improved school attendance (Little & Harris, 2003). However, we have found that the pooled effects on standardized test scores from multiple experimental studies do not show positive impacts. Our meta-analysis of rigorous experimental design studies arguably presents a more credible understanding of overall program impacts than the findings from any single study.

Overall, there are multiple areas of disagreement between the findings of this review and other prior reviews in this field. Although it is not possible to know whether the conflicting findings are a function of differences between the program models, the study methodologies, or the analytic techniques, this review does suggest that initial optimism about program impacts are tempered when tighter and more stringent review criteria are applied and meta-analytic techniques pool program impacts across multiple studies.

### **Discussion**

The U.S. Department of Education is moving swiftly towards a reliance on evidence from rigorous studies, such as those included in this review, when making policy and programming decisions. While this review has included the most rigorous studies conducted of after-school programs that are currently of great policy interest due to their inclusion of academic support components, reviewers note that the collected evidence is not sufficient to make any policy or programming recommendations. While some areas of promise do exist—supervision and participation in activities—these pooled impacts need to be tested with further research.

Similarly, further rigorous research may verify the extent to which similar programs can have much significant impact on the measured outcomes. The primarily null impacts (84 percent) reported by the studies in this review could be a function of limited duration of the intervention (five to nine months) or the relatively low participation rates across the studies. For example, Lauer (2002) reported that only 15 percent of those youth enrolled in the Cooke Middle School After-School Recreation program attended more than half of the scheduled days. Only 38

Table 26: Comparison of Findings Between this Review and Prior Reviews

Review	Study methodologies	Definition of intervention	Study overlap with Zief & Lauver	Findings	
				Prior review	Zief and Lauver <sup>25</sup>
Eccles and Templeton (2002)	RFTs, QEDs, & NEx (separated NEx out for analyses)	Extracurricular, sports, and leisure activities	None	Positive impacts for promoting positive behaviors and reducing negative behaviors	Disagree
Fashola (1998)	RFTs, QEDs, & NEx	Educational focus with evidence or promise of effectiveness	None	Inconclusive results due to few rigorous studies at time of review	Cannot agree or disagree
Hollister (2003)	RFTs	Out-of-school time programs that focus on positive youth development	None	Mentoring/tutoring models promote positive in-school and out-of-school outcomes; parent involvement and life skills training curricula positively impact out-of-school outcomes	Cannot agree or disagree
Kane (2004)	RFTs and QEDs	Primarily academic programs with recreation and youth development programming	U.S. Department of Education, 2003	Parents more likely to participate in schools, students more likely to complete homework, inconsistent but encouraging GPA outcomes	Disagree on homework and parent participation; agree on grades
Little and Harris (2003)	RFTs and QEDs	Academic, prevention, and youth development programs	U.S. Department of Education, 2003 Weisman et al., 2001	Significant improvements in multiple outcome areas	Disagree on most; agree on grades and new experiences
Scott-Little et al. (2002)	RFTs, QEDs, & NEx (separated NEx out for analyses)	Wide range of after-school program models	None	Positive impacts on standardized test scores and homework completion; RFTs suggest improved social/emotional outcomes	Disagree

Note: RFT = randomized field trial; QED = quasi-experimental design; NEx = Non-Experimental

percent of the elementary 21<sup>st</sup> CCLC students participated in the program more than half of the days it was offered (U.S. Department of Education, 2003). In a recent review, Kane (2004) also draws attention to the sporadic attendance at programs that operate on a voluntary, drop-in basis.

Alternatively, the null findings could also suggest that after-school programs similar to those included in this review (i.e., programs that offer a range of academic, recreational, and enrichment activities) may not be an effective means to achieve the outcomes that after-school

<sup>25</sup>We limit the contributed findings from our review to the pooled effects of two or more studies calculated using meta-analytic techniques.

programs in theory hold such promise to impact. Students' experiences outside of such programs could have more of an impact on the measured outcomes. Kane (2004) introduces a similar argument, suggesting that significant impacts on academic outcomes cannot be achieved with an intervention of such limited dosage, especially considering, for example, that a full year of sixth grade classroom instruction is associated with only a .25 to .30 standard deviation gain in reading and math achievement.

The primary product of these review efforts can instead be a guide for the conduct of a systematic review and analysis of like programs and evaluations, and a benchmark from which future after-school program review efforts can begin. Furthermore, this review has served to identify more clearly the knowledge that is lacking and how we might acquire it. Below, we suggest some answers.

1. Rigorously test alternative models and target populations.  
We believe that there is significantly more variation among program models that would fall under the heading of a traditional after-school program with academic support components. These models need to be identified and tested using experimental methods.
2. Promote replication of experimental design studies.  
Future experimental design evaluations need consistency in their measured outcomes to promote greater use of meta-analytic techniques to measure overall program impacts. This could be accomplished by employing a common research design template to multiple, small-scale, mixed-methods studies.
3. Extend data collection.  
The studies in this review collected data after five to nine months of youth exposure to the programs. What are the longer term impacts of after-school programs? Besides the national study of 21<sup>st</sup> CCLCs, we know of no other longitudinal experimental design study currently being conducted in the field.
4. Consider additional outcome areas.  
In this era of welfare reform, it might be important to understand how this type of intervention targeting low-income youth may impact parents' job attendance and retention or parental levels of stress while balancing the demands of work and child-care. Such parental outcomes could arguably mediate student social and emotional, behavioral, and academic outcomes.
5. Implement complementary process evaluations.  
The detailed descriptions of program operations and processes of the Cooke Middle School After-School Recreation Program evaluation design were quite valuable when interpreting the impact estimates (Lauver, 2002). If more experimental evaluations had such design components, the results of future meta-analyses could be more helpful to policymakers, program developers, and practitioners to guide the design and delivery of programs.

6. Improve study reporting.

No study provided all of the data we needed to judge the rigor of implementation or to the prepare data for meta-analysis. We recommend that researchers become more transparent and comprehensive in their study reporting, perhaps following a template similar to the checklist we used to judge study quality (Appendix 3).

## References

References marked with an asterisk indicate prior reviews searched for relevant studies.

- \*Afterschool Alliance. (2003). *Afterschool Alliance backgrounder: Formal evaluations of afterschool programs*. Washington, DC: Afterschool Alliance. Retrieved from [www.afterschoolalliance.org/issuebr.cfm](http://www.afterschoolalliance.org/issuebr.cfm).
- Agodini, R. & Dynarski, M. (2001). *Are experiments the only option? A look at dropout prevention programs*. Princeton, NJ: Mathematica Policy Research, Inc.
- Baker, D., & Witt, P. A. (1996). Evaluation of the impact of two after-school recreation programs. *Journal of Park and Recreation Administration*, 14(3), 23-44.
- \*Beckett, M., Hawken, A., & Jackowitz, A. (2001). *Accountability for after-school care: Devising standards and measuring adherence to them*. Santa Monica, CA: RAND.
- Belden Russonello & Stewart Research and Communications. (2001). *Principals and after-school programs: A survey of preK-8 principals*. Washington, DC: National Association of Elementary School Principals.
- \*Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., & Hawkins, J. D. (1998). *Positive youth development in the United States: Research findings on evaluations of positive youth development programs*. Seattle, WA: University of Washington, Social Development Research Group.
- Cooper, H., & Hedges, L. (Eds.) (1994). *The Handbook of Research Synthesis*. New York, NY: Russell Sage Foundation.
- De Kanter, A. (2001). After-school programs for adolescents. *NASSP Bulletin*, 85(626), 12–21.
- Dwyer, K. M., Richardson, J. S., Danley, K. L., Nansen, W. B., Sussman, W. Y., Brannon, B., Det, C. W., Johnson, C. A., & Flay, B. R. (1990). Characteristics of eighth-grade students who initiate self-care in elementary and junior high school. *Pediatrics*, 86, 448-454.
- Dynarski, M., James-Burdumy, S., Mansfield, W., Mayer, D., Moore, M., Mullens, J., & Silva, T. (2001). *A broader view: The national evaluation of the 21<sup>st</sup> Century Community Learning Centers program*. Princeton, NJ: Mathematica Policy Research, Inc.
- \* Eccles, J. S., & Templeton, J. (2002). Extracurricular and other after-school activities for youth. In W.S. Secada (Ed.), *Review of Research in Education*, 26, 113–180. Washington, DC: American Education Research Journal.

- \*Fashola, O. S. (1998). *Review of extended-day and after-school programs and their effectiveness*. Baltimore, MD: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk (CRESPAR).
- Foley, E. M, Eddins, G., & Fenton, R. (2000). *Impact of the Virtual Y on children's classroom behavior and school attendance*. New York: Fordham University, National Center for Schools and Communities.
- Glazerman, S., Levy, D., & Myers, D. (2003). *Nonexperimental replications of social experiments: A systematic review*. Princeton, NJ: Mathematica Policy Research, Inc.
- Grossman, J. B., Price, M., Fellerath, V., Jucovy, L., Kotloff, L., Raley, R., & Walker, K. (2002). *Multiple choices after school: Findings from the Extended-Service School Initiative*. Philadelphia, PA: Public/Private Ventures.
- Guyat, G. H., DiCenso, A., Farewell, V., Willan, A., & Griffith, L. (2000). Randomized trials versus observational studies in adolescent pregnancy prevention. *Journal of Clinical Epidemiology*, 53, 167-174.
- Hahn, A., Leavitt, T., & Aaron, P. (1994). *Evaluation of the Quantum Opportunities Program: Did the program work?* Waltham, MA: Brandeis University.
- \*Harvard Family Research Project. (2003). *Out-of-school time evaluation database*. Retrieved from <http://gseweb.harvard.edu/~hfrp/projects/afterschool/evaldatabase.html>
- Hofferth, S. L. (1995). Out-of-school time: Risk and opportunity. In T.R. Swartz & K.M. Weigert (Eds.). *America's working poor*. Notre Dame, IN: University of Notre Dame Press.
- \*Hollister, R. (2003). *The growth in after-school programs and their impact*. Washington, DC: The Brookings Institution.
- Huang, D., Gribbons, B., Kim, K. S., Lee, C., & Baker, E. (2000). *A decade of results: The impact of LA's BEST after school enrichment program on subsequent student achievement and performance*. Los Angeles, CA: UCLA Center for the Study of Evaluation (CSE).
- \*James, D.W. (Ed.). (Unknown date). *MORE things that DO make a difference for youth*. Washington, DC: American Youth Policy Forum.
- \*James, D.W., & Donahue, C. (Eds.). (Unknown date). *SOME things do make a difference for youth*. Washington, DC: American Youth Policy Forum.
- Jones, M. B., & Offord, D.R. (1989). Reduction of antisocial behavior in poor children by nonschool skill-development. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 30(5), 737-750.

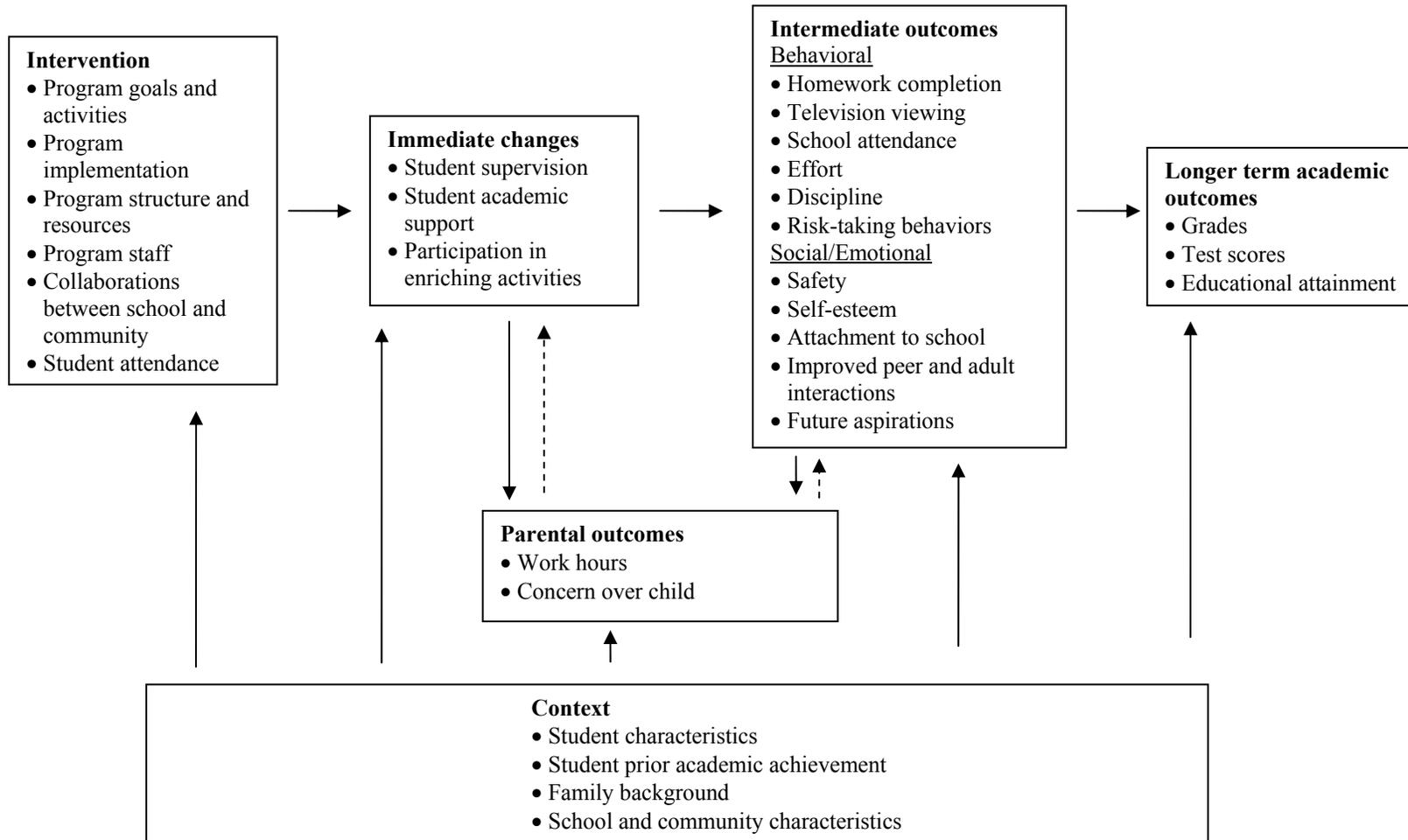
- \*Kane, T. J. (2004). *The impact of after-school programs: Interpreting the results of four recent evaluations*. Working paper. New York, NY: W.T. Grant Foundation.
- \*Kirby, D. (1997). *No easy answers: Research findings on programs to reduce teen pregnancy*. Washington, DC: National Campaign to Prevent Teen Pregnancy.
- \*Kirby, D. (2001). *Emerging answers: Research findings on programs to reduce teen pregnancy*. Washington, DC: National Campaign to Prevent Teen Pregnancy.
- Lauver, S. (2002). *Assessing the benefits of an after-school program for urban youth: Results of an experimental design and process evaluation*. Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia, PA.
- Le, V.-N., and Hamilton, L.S. (2001). *Achievement gains in math and reading by participants of the Foundations after-school enrichment program* (PM-1265-EDU). Santa Monica, CA: RAND.
- Lipsey, M.W., & Wilson, D.B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications, Inc.
- \*Little, P., & Harris, E. (2003). *A review of out-of-school time program quasi-experimental and experimental evaluation results*. Out-of-School Time Evaluation Snapshot: Number 1. Boston, MA: Harvard Family Research Project.
- Long, S., & Clark, S. (1998). *The new child care block grant: State funding choices and their implications*. Washington, DC: Urban Institute.
- LoSciuto, L., Freeman, M., Harrington, E., Altman, B., & Lanphear, A. (1997). An outcome evaluation of the Woodrock Youth Development Project. *Journal of Early Adolescence*, 17(1), 51-66.
- LoSciuto, L., Hilbert, S. M., Fox, M. M., Porcellini, L., & Lanphear, A. (1999). A two-year evaluation of the Woodrock Youth Development Project. *Journal of Early Adolescence*, 19(4), 488-507.
- McLaughlin, M., & Irby, M. (1994). Urban sanctuaries: Neighborhood organizations that keep hope alive. *Phi Delta Kappan*, 76(4), 300-306.
- National Institute on Out-of-School Time. (2003). *Fact sheet on school-age children's out of school time*. Retrieved from <http://www.wellesley.edu/WCW/CRW/SAC/factsheet.pdf>.
- \*National Research Council and Institute of Medicine. (2002). *Community programs to promote youth development*. Committee on Community-Level Programs for Youth. J. Eccles & J. A. Gootman (Eds.). Board on Children, Youth, and Families, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

- Newman, S. A., Fox, J. A., Flynn, E. A., & Christeson, W. (2000). *America's after-school choice: The prime time for juvenile crime or youth enrichment and achievement*. Washington, DC: Fight Crime: Invest in Kids.
- Osofsky, J. D. (1999). The impact of violence on children. *The Future of Children, 9*. Los Altos, CA: The David and Lucille Packard Foundation.
- Phillips, R. (1999). Intervention with siblings of children with developmental disabilities from economically disadvantaged families. *Families in Society, 80*(6), 569-577.
- Posner, J. K., & Vandell, D.L. (1994). Low-income children's after-school care: Are there beneficial effects of after-school programs? *Child Development, 65*, 440-456.
- Posner, J. K., & Vandell, D. L. (1999). After-school activities and the development of low-income urban children: A longitudinal study. *Developmental Psychology, 35*, 868-879.
- Proscio, T., & Whiting, B. (2004). *After-school grows up: How four large American cities approach scale and quality in after-school programs*. New York: The After School Project.
- Richardson, J. L., Dwyer, K., McGuigan, K., Hansen, W., Dent, C., Johnson, C. A., Sussman, S., Brannon, B., & Flay, B. (1989). Substance use among eighth-grade students who take care of themselves after school. *Pediatrics, 84*, 556-566.
- Ross, J. G., Saavedra, P. H., Shur, G. H., Winter, F., & Felner, R. D. (1992). The effectiveness of an after-school program for primary grade latchkey students on precursors of substance abuse. *Journal of Community Psychology, OSAP Special Issue*.
- \*Roth, J., Brooks-Gunn, J., Murray, L., & Foster, W. (1998). Promoting healthy adolescents: Synthesis of youth development program evaluations. *Journal of Research on Adolescence, 8*(4), 423-459.
- Schinke, S. P., Cole, K. C., & Poulin, S. R. (2000). Enhancing the educational achievement of at-risk youth. *Prevention Science, 1*(1), 51-60.
- \*Scott-Little, C., Hamann, M. S., & Jurs, S. G. (2002). Evaluations of after-school programs: A meta-evaluation of methodologies and narrative synthesis of findings. *American Journal of Evaluation, 23*(4), 387-419.
- Tucker, C. M., Herman, K. C., Reid, A. D., Keefer, N. L., & Vogel, D. L. (1998). The research-based model partnership education program: A 4-year outcome study. *Journal of Research and Development in Education, 32*(1), 32-37.

- U.S. Bureau of Labor Statistics. (2000). *Families with own children: Employment status of parents by age of youngest child and family type, 2001-2002 annual averages*. Retrieved from <http://www.bls.gov/news.release/famee.t04.htm>.
- U.S. Department of Education & U.S. Department of Justice. (2000). *Working for children and families: Safe and smart after-school programs*. Washington, DC: Government Printing Office.
- U.S. Department of Education, Office of the Under Secretary. (2003) *When schools stay open late: The national evaluation of the 21<sup>st</sup> Century Community Learning Centers program, first year findings*. Washington, DC: Author.
- U.S. Department of Health and Human Services. (1995). *Adolescent time use, risky behavior, and outcomes: An analysis of national data*. Retrieved from <http://aspe.hhs.gov/hsp/cyp/xstimuse.htm>.
- Vandivere, S., Tout, K., Zaslow, M., Calkins, J., & Capizzano, J. (2003). *Unsupervised time : Family and child factors associated with self-care*. Washington, DC: Urban Institute.
- Washburn, R. A. (2004). *Prevention of obesity in YMCA daycare centers*. Lawrence, KS: University of Kansas, Lawrence.
- Weisburd, D., Lum, C., & Petrosino, A. (2001). Does research design affect study outcomes in criminal justice? *The Annals, AAPSS*, 578, 50-70.
- Weisman, S. A., Soule, D. A., Womer, S. C., & Gottfredson, D. C. (2001). *Maryland After School Community Grant Program: Report on the 1999-2000 school year evaluation of the Phase I after-school programs*. College Park, MD: University of Maryland. (Experimental evaluation of Programs 4, 17 and 21.)
- Welsh, M. E., Russell, C. A., Williams, I., Reisner, E. R., & White, R. N. (2002). *Promoting learning and school attendance through after-school programs: Student-level changes in educational performance across TASC's first three years*. Washington, DC: Policy Studies Associates.
- Wilson, D., & Lipsey, M. (2001). The role of method in treatment effectiveness research: Evidence from meta-analysis. *Psychological Methods*, 6(4), 413-429.

**APPENDIX 1**  
**LOGIC MODEL**

## Logic Model for Understanding the Theory of Change for Low-Income Elementary Youth in an After-School Program



Note: Adapted from Dynarski et al. (2001) and Lauer (2002).

## **APPENDIX 2**

### **DATABASE SEARCH STRATEGY**

## Database Search Strategy

### I. Database search strategy for ERIC, Dissertation Abstracts, and Education Index

Search Engine = Ovid Technologies, Inc., Time Period = 1982-2002, Language = English

---

- 1 exp After School Centers
- 2 limit 1 to yr=1982-2002
- 3 exp After School Programs
- 4 limit 3 to yr=1982-2002
- 5 exp After School Education
- 6 limit 5 to yr=1982-2002
- 7 2 and 4 and 6
- 8 2 or 4 or 6
- 9 Evaluation.mp.
- 10 Limit 9 to yr=1982-2002
- 11 Outcome.mp.
- 12 Limit 11 to yr=1982-2002
- 13 Impact.mp.
- 14 Limit 13 to yr=1982-2002
- 15 10 or 12 or 14
- 16 8 and 15

### II. Database search strategy for PsychINFO

We encountered difficulties executing the above search in PsychINFO using the Ovid system at the University of Pennsylvania because “after” was considered a stop term and the search was terminated. An Ovid Silverplatter reference librarian conducted the following search at the request of the University of Pennsylvania library staff:

- 1 after school center? Or after-school center? Or afterschool center?
- 2 after school program? Or after-school program? Or afterschool program?
- 3 afterschool or after-school or (after adj school)
- 4 #1 or #2

We recommend that future searches be conducted after consultation with a reference librarian, such as the Campbell Collaboration’s Information Retrieval Methods Group Liaison, to assist in the refinement of the above search strategies.

**APPENDIX 3**

**CHECKLIST FOR JUDGING STUDY QUALITY**

## Checklist for Judging Study Quality

Section and Topic	Item Number	Descriptor	Comments
<b>INTRODUCTION</b>			
Title and abstract	1	Study design	
Background	2	Relationship of evaluator to intervention	
	3	Relationship of study sponsor to intervention	
	4	Explanation of the rationale for the intervention	
Objectives	5	Specific goals/objectives and hypotheses	
	6	Logic model or theory of change	
<b>METHODS</b>			
Participants	7	Eligibility criteria for participants (i.e., target population)	
	8	Explanation of recruitment procedures	
Intervention	9	Precise details of the intended intervention	
	10	Precise details on the implementation of the intervention	
	11	Information about the activities of the control group	
	12	Information on possible contamination	
Outcomes	13	Clearly defined primary and secondary outcome measures	
	14	Outcome measures aligned with the goals of the intervention	
	15	Explanation of measurement instruments and information regarding their validity and reliability	
	16	Methods used to enhance the quality of the data (supplemental studies, multiple evaluations, training of data collectors)	
Sample size	17	Size of treatment and control groups	
	18	Use of power analysis to determine sample size	
Randomization sequence	19	Explanation of the method used to generate the random allocation sequence, including details of any restrictions (e.g., blocking, stratification)	
	20	Parental consent for study participation received prior to random assignment	

Section and Topic	Item Number	Descriptor	Comments
Statistical methods	21	Statistical methods used to compare groups for primary outcome(s) and for additional analyses, such as subgroup analyses	
	22	Appropriateness of methods chosen	
	23	Pre-intervention measures of outcomes and other important variables collected at baseline and incorporated into the analysis	
<b>RESULTS</b>			
Attrition	24	Number in each group who withdrew from study	
	25	Number in each group who were lost to followup	
	26	Number excluded from analysis (give reason)	
	27	Attrition > 20%: Completers statistically compared to non-completers	
	28	Attrition > 20%: Baseline equivalence of analytic sample demonstrated	
Intention-to-treat	29	Whether the analysis was by “intention-to-treat”	
Outcomes and data reporting	30	For each outcome, a summary of results per group	
	31	Means and SDs reported	
	32	<i>p</i> -values and degrees of freedom reported	
	33	Effect sizes reported	
	34	Other value reported (specify)	
<b>CONCLUSIONS</b>			
Interpretation	35	Interpretation of the results, taking into account study hypotheses and sources of potential bias or imprecision.	
	36	Use of observational data to understand impact results	
External validity	37	Generalizability of results	
	38	Replicability of intervention	
Overall evidence	39	General interpretation of the results in the context of current evidence.	

Note: Adapted from the *What Works Clearinghouse Study Review Standards* (U.S. Department of Education, 2005).



Information deemed critical for inclusion in review.

**APPENDIX 4**

**FULL DESCRIPTIONS OF STUDIES INCLUDED IN REVIEW**

**21<sup>st</sup> Century Community Learning Centers**

U.S. Department of Education, Office of the Undersecretary. (2003). *When schools stay open late: The national evaluation of the 21<sup>st</sup> Century Community Learning Centers program, first year findings*. Washington, DC: Author.

**Program description**

Funding source Program location Program operator	Program times Staffing Links to school day	Target Population Demographics	Program goals and components	
			Goals	Offerings
Federal (U.S. Department of Education)  18 mostly urban public elementary schools (7 sites provided impact estimates)  Grant funds directly to districts and schools; program operators more varied  Collaborations between local organizations common	Immediately after school, some Saturdays and summers  School staff, paraprofessionals, older students  Few links to school day at most sites	Low-income, low-achieving minority (mostly African American) students  Grades: K–6  46% male 67% African American, 19% Latino, 10% white	<ul style="list-style-type: none"> <li>• Improve academic outcomes</li> <li>• Provide a safe after-school environment where children can develop socially, engage in recreational activities, and participate in cultural events</li> </ul>	Homework assistance; reading, math and science supplemental curricular activities; state or district test preparation; recreational activities, clubs, recess, and snack

Note: There was some variation between the programs' operations and activities included in the impact analysis. The table data represent the operations and activities that were more common across the centers.

**21st Century Community Learning Centers (continued)**

<b>Evaluation</b>			
Evaluator Time period Level of RA Baseline sample size Response rate Analysis	Outcomes and Findings		Implications
	Measures	Results	
Mathematica Policy Research, Inc.  Fall 2000–spring 2001  Individual  Baseline $N = 968$ ( $TxN = 587, CoN = 381$ )  94 percent survey response rate  OLS and logistic regression	Student questionnaires, school records, parent questionnaires, teacher questionnaires  Outcomes included academic performance and homework completion, behavior, feelings of safety, and personal and social development.	Significant ( $p < .05$ ): Participants had higher social studies grades, were more likely to have parents who were involved with homework or school events, more likely to be in non-parent adult care, less likely to be in their own home, more likely to be at school (or other place) for after-school activities, stayed after school more days per week, less likely to be involved in organized clubs (B & G Clubs, Scouts), and more likely to take care of brother or sister. Several additional outcomes (contextual and behavioral) approach significance ( $.05 < p < .10$ ).	More frequent attendees did not have better outcomes, and programs with greater emphasis on academics did not show greater academic outcomes. Participants attended programs for an average of 1.8 days—attention needs to be paid to attracting and retaining student interest and to the content of the academic programming. Eighteen study sites may not be a representative sample—participated because of greater demand for program slots than availability at these sites and willingness to be a part of the random assignment process.

**Maryland After-School Community Grant Program: Program 21**

Weisman, S. A., Soule, D. A., Womer, S. C., & Gottfredson, D. C. (2001). *Maryland After-School Community Grant Program: Report on the 1999–2000 school year evaluation of the phase 1 after-school programs*. College Park, MD: University of Maryland.

<b>Program description</b>				
Funding source Program location Program operator	Program times Staffing Links to school day	Target population Demographics	Program goals and components	
			Goals	Program activities
Maryland Governor’s Office of Crime Control and Prevention with federal funds from the U.S. Dept. of Justice  Baltimore, MD elementary school  School	Tues.–Fri., 3:30–6 p.m.  School teachers, principal, guidance counselor, police officer, and student volunteers	Fourth- and fifth-grade students, “low to middle class,” predominantly Caucasian student population	Two major goals were to significantly lower levels of self-reported delinquency and to significantly lower levels of self-reported drug use. Other objectives included (a) decreasing the number of unsupervised after-school hours; (b) increasing bonding to school, community, and family; (c) increasing academic performance; (d) decreasing positive attitudes towards substance use and illegal behaviors; (e) decreasing negative peer influence; and (f) improving social skills.	A balanced program of academic and social skills instruction with additional recreation activities. It was modeled after the Police Athletic League programs and the Prince George’s County Latchkey program. Recreational activities include board and video games, pool and ping-pong, and structured sports.

**Maryland After-School Community Grant Program: Program 21 (continued)**

<b>Evaluation</b>			
Evaluator Time period Level of RA Baseline sample size Response rate Analysis	Outcomes and findings		Implications
	Measures	Results	
University of Maryland  1999–2000  Individual  Baseline $N = 49$ ( $TxN = 22$ , $CoN = 27$ )  94 percent survey response rate  Comparison of means (t-test)	What About You? Survey (Gottfredson, 1991)  Items from the Social Skills Rating System (SSRS) elementary form (Gresham & Elliott, 1990)  Communities That Care Youth Survey (Arthur et al., 1995)  School records (GPA)	There were mostly negative program effects on participants, although only one was statistically significant—participants were less likely to be under the influence of positive peers ( $p < .05$ ). However, there were moderate (nonsignificant) positive effects found for decreasing participants' negative peer influences.	This program could greatly benefit from a more thorough implementation study. In particular, it would be helpful to understand the peer culture in the program group and how students are influencing one another. It would also be beneficial to know if the control group are participating in other after-school activities that are having some influence on these results. This study demonstrates the importance of assessing program quality and its effect on program outcomes.

**Maryland After-School Community Grant Program: Program 17**

Weisman, S. A., Soule, D. A. Womer, S. C., & Gottfredson, D. C. (2001, June). *Maryland After-School Community Grant Program: Report on the 1999–2000 school year evaluation of the phase 1 after-school programs*. College Park, MD: University of Maryland.

<b>Program description</b>				
Funding source Program location Program operator	Program times Staffing Links to school ay	Target population Demographics	Program Goals and Components	
			Goals	Program activities
<p>Maryland Governor’s Office of Crime Control and Prevention with federal funds from the U.S. Dept. of Justice</p> <p>Community center in Baltimore, MD</p> <p>Community-based organization</p>	<p>Tues.–Fri., 3:30–6 p.m.</p> <p>Managerial staff (with college degrees) of the Foundation</p> <p>Individualized academic assessments and plans made for each child; however, no other direct links to school day</p>	<p>20 fourth- and fifth-grade students at an inner-city school just north of Baltimore’s Inner Harbor</p> <p>100% minority</p>	<p>Two major goals were to significantly lower levels of self-reported delinquency and to significantly lower levels of self-reported drug use. Other objectives included (a) decreasing the number of unsupervised after-school hours; (b) increasing bonding to school, community, and family; (c) increasing academic performance; (d) decreasing positive attitudes towards substance use and illegal behaviors; (e) decreasing negative peer influence; and (f) improving social skills.</p>	<p>A combination of tutoring and homework assistance, social skills instruction, and recreational opportunities including sailing and rowing, Chesapeake Bay and environmental studies, service learning, and monthly field trips</p>

**Maryland After-School Community Grant Program: Program 17 (continued)**

<b>Evaluation</b>			
Evaluator Time period Level of RA Baseline sample size Response rate Analysis	Outcomes and findings		Implications
	Measures	Results	
University of Maryland  1999–2000  Individual  Baseline $N = 44$ ( $TxN = 19$ , $CoN = 25$ )  82 percent survey response rate  Comparison of means (t-test)	What About You? Survey (Gottfredson, 1991)  Items from the Social Skills Rating System (SSRS) elementary form (Gresham & Elliott, 1990)  Communities That Care Youth Survey (Arthur et al., 1995)  School records (GPA)	Program participants spent significantly less time in self-care ( $p < .05$ ). There was also a significant positive effect on time spent with a positive peer group ( $p < .05$ ). Moderate positive (but not significant) effects were found for parental supervision and positive attitudes towards substance use. A moderate negative nonsignificant effect was found for improvement in social skills. There were no differences in academic performance between participants and those students in the control group.	This program offers unique learning opportunities to participants. This program would benefit from a more intensive, longitudinal study evaluating long-term effects on academic achievement and other skills.

**Maryland After-School Community Grant Program: Program 4**

Weisman, S. A., Soule, D. A. Womer, S. C., & Gottfredson, D. C. (2001, June). *Maryland After-School Community Grant Program: Report on the 1999–2000 school year evaluation of the phase 1 after-school programs.* College Park, MD: University of Maryland.

Program description				
Funding source Program location Program operator	Program times Staffing Links to school day	Target population Demographics	Program goals and components	
			Goals	Program activities
Maryland Governor’s Office of Crime Control and Prevention with federal funds from the U.S. Dept. of Justice  Beach Elementary School, Chesapeake Beach, MD  School	Mon.–Thurs., 4–6 p.m.  School employees—teachers, aides, librarian, nurse, principal  Individualized academic plans made; school-day teachers were also the after-school staff members	Fourth- and fifth-grade students  Demographics unknown	Two major goals were to significantly lower levels of self-reported delinquency and to significantly lower levels of self-reported drug use. Other objectives included (a) decreasing the number of unsupervised after school hours; (b) increasing bonding to school, community, and family; (c) increasing academic performance; (d) decreasing positive attitudes towards substance use and illegal behaviors; (e) decreasing negative peer influence; and (f) improving social skills.	Academic achievement, social skills instruction and practice, and recreational activities including arts and crafts, board games, computer use, and basketball. All participants were members of the Just Say No Club to teach drug awareness and prevention. Field trips to the Naval Academy and Baltimore were also offered.

**Maryland After-School Community Grant Program: Program 4 (continued)**

<b>Evaluation</b>			
Evaluator Time period Level of RA Baseline sample size Response rate Analysis	Outcomes and findings		Implications
	Measures	Results	
University of Maryland  1999–2000  Individual  Baseline $N = 63$ ( $TxN = 32$ , $CoN = 31$ )  92 percent survey response rate  Comparison of means (t-test)	What About You? Survey (Gottfredson, 1991)  Items from the Social Skills Rating System (SSRS) elementary form (Gresham & Elliott, 1990)  Communities That Care Youth Survey (Arthur et al., 1995)  School records (GPA)	Program reduced hours in self-care ( $p < .05$ ) and increased involvement in constructive activities ( $.05 < p < .10$ ). The program had moderate negative (nonsignificant) effects on academic performance and attachment to prosocial adults.	This program appears to be a fairly traditional after-school program offering supervision, homework assistance, and recreation for students. Participation in the program is moderate, and moderate effect sizes were found in a number of the outcome areas. A longitudinal evaluation of the program is suggested.

**Cooke Middle School After-School Recreation Program, Philadelphia, PA**

Lauver, S.C. (2002). *Assessing the benefits of an after-school program for urban youth: An impact and process evaluation*. Unpublished doctoral dissertation, University of Pennsylvania.

<b>Program description</b>				
Funding source Program location Program operator	Program times Staffing Links to school day	Target population Demographics	Program goals and components	
			Goals	Program activities
School district and federal 21 <sup>st</sup> CCLC funds  Urban middle school in Philadelphia, PA  School	Mon.–Thurs., 5–7 p.m.  Teachers, college student volunteers  Program staffed by school-day teachers	Fifth–eighth grade  85% African American, 12% biracial, 3% other minority  36% male, 64% female  8% English language learners  85% of school eligible for free/reduced lunch	To promote students’ physical, emotional, and social well-being. Major objectives are the following: <ul style="list-style-type: none"> <li>• Providing a safe, structured, adult-supervised environment that encourages positive adolescent development</li> <li>• Providing opportunities for exercise, fitness skills, and wellness</li> <li>• Extracurricular activities that encourage students’ social, emotional, intellectual, and physical skills</li> <li>• Teaching children how to get along in a nonthreatening environment, where they interact with teachers outside of the normal educational setting</li> </ul>	Dance program, computer games, basketball, jump rope, fitness center, board games, quiet space for reading/homework, and arts and crafts

**Cooke Middle School After School Recreation Program, Philadelphia, PA (continued)**

<b>Evaluation</b>			
Evaluator Time period Level of RA Baseline sample size Response rate Analysis	Outcomes and findings		Implications
	Measures	Results	
University of Pennsylvania  January–June, 2000  Individual  Baseline $N = 222$ ( $TxN = 124$ , $CoN = 98$ )  99 percent survey response rate  OLS and logistic regression controlling for preprogram differences between groups	Hours per week participating in (a) arts, music, or dance; (b) strength-training exercises in a fitness center or home; and (c) exercising or playing sports  <i>Hours per day (a) watching TV, (b) in self-care, and (c) doing homework</i>  Educational aspirations, attendance, academic grades and GPA, and student behavior	The program had a significant and positive effect on program participants' time spent doing homework and on their educational aspirations. Eighty-nine percent of program participants spent one or more hours per week on homework, as compared to 76% of the students in the control group. Furthermore, 94% of program participants hoped to attend college, versus 85% of students in the control group.  Significantly more program than control group youth spent one or more hours per week working out in a fitness center (80% versus 61%, respectively). There were no measurable benefits of the after-school program on other extracurricular activities, academic grades or standardized test scores, in-school behavior or attendance, or on time spent watching television or in self-care. Students in both groups spent approximately four hours watching television daily, and 65% spent one or more hours in self-care daily.	The program had relatively poor student attendance because it was perceived as a “drop-in” program by participants. This limitation may have diluted the potential impact of the program. The program also suffered from high staff turnover, which is common in many after-school programs. Recommendations for program design were to strengthen staff recruitment and retention, enhance program activities, promote regular student attendance, incorporate a homework assistance program, negotiate access to greater school space, and provide transportation.

