Harvard Family Research Project’s (HFRP) Issues and Opportunities in Out-of-School Time Evaluation briefs highlight current research and evaluation work in the out-of-school time field. These documents draw on HFRP’s research work in out-of-school time in order to provide practitioners, funders, evaluators, and policymakers with information to help them in their work. This latest brief reviews developmental research and out-of-school time program evaluations to examine three research-based indicators of attendance—intensity, duration, and breadth—offering different models for how attendance in out-of-school time programs can influence youth outcomes.

A growing evidence base suggests that participation in out-of-school time (OST) programs can make a positive difference in the lives of young people. Researchers and practitioners assert that high quality, organized OST activities have the potential to support and promote youth development because they (a) situate youth in safe environments; (b) prevent youth from engaging in delinquent activities; (c) teach youth general and specific skills, beliefs, and behaviors; and (d) provide opportunities for youth to develop relationships with peers and mentors. In fact, evidence increasingly shows that youth participation in quality OST activities influences their current outcomes, which in turn impact outcomes into adulthood. Participation in OST activities is predictive of academic success as measured through test scores, absenteeism, school dropout rates, homework completion, school grades, and course enrollment. Further, some suggest that OST programs can provide the opportunity to develop critical “21st century” skills that include problem solving and interpersonal and communication skills, as well as proficiency in the “basics.”

Participation in OST activities is also related to multiple indicators of positive social development. Research shows that participation is related to more prosocial and less aggressive behavior with peers, multiple aspects of friendships, and lower feelings of depression and problem or delinquent behavior. While these studies examine important relations between overall participation and outcomes, a key question remains unanswered: How much participation, in what kinds of programs, and for which participants is necessary to improve outcomes for youth? Implicit in this overarching question is the need for accurate and meaningful ways to assess youth participation in OST programs.

The first part of this issue brief draws on developmental research and OST program evaluations to examine three research-based indicators of participation—intensity of attendance, duration of attendance, and breadth of attendance. The subsequent sections summarize research that links overall participation to outcomes, the research organized using the three indicators. Finally, the brief offers three models that help explain the relationship between attendance and outcomes.

Before we continue, it is important to make a distinction between the terms participation and attendance. Participation is defined as active involvement in an after school program. Attendance, which can be measured...
in a variety of ways (daily, weekly, by activity, etc.), is generally an indication of the time youth spend in activities. To date, many researchers and evaluators interested in measuring youth participation have grouped youth into one of two categories: those who attend OST activities and those who do not. Although these groupings have been and will continue to be useful in our understanding of overall OST program participation, measuring attendance in such global terms glosses over critical information about how often youth attend activities, how many years they attend, and whether they participate in one activity or several.

Beyond the scope of this brief is a presentation of information about involvement and engagement as important components of overall participation. Attendance is a necessary but not sufficient indicator of participation; assessing youth participation and its links to outcomes requires the creation of multiple indicators that capture information about involvement and engagement, as well as attendance. For articles that address how involvement and engagement influence attendance, see Youth Engagement Resources at the end of this brief.

A Note on Our Methodology
A description of the methodology used in searching for and selecting studies for this issue brief can be found in Appendix A. Briefly, published and unpublished papers, including program evaluations, were identified through several searches. We found 83 studies that met all of the methodological and design criteria. From this group, 27 studies included findings relevant to the issues discussed in this brief. These studies are listed in Appendix B. The studies cover a range of youth development indicators, OST program and participation indicators, and evaluation and academic research.

Indicators of Attendance
Attendance in OST programs is not solely a yes or no construct. Attendance, or the time youth spend in OST activities, can be measured by four indicators:

1. Whether youth spend any time in an OST activity or program
2. The intensity of their attendance
3. The duration of their attendance
4. The breadth of their attendance

All of these indicators capture unique dimensions of overall participation. Yet, to date, the first indicator—whether or not youth spend any time in an OST activity—has been the most frequently used. Nearly 70% of the studies identified for this brief used only this indicator, which compares youth who spend any time in an activity with youth who do not participate at all. A likely explanation is that, of the various indicators, this one measuring absolute attendance is often the most cost effective and easy to measure. However, it is also the indicator that captures the least amount of information concerning participation. Intensity, duration, and breadth, the more nuanced indicators, can yield a more complete understanding of how attendance influences outcomes. Each is defined below.

Intensity. Intensity is the amount of time youth attend a program during a given period. Intensity has been measured in terms of hours per day, days per week, and weeks per year; it varies across programs and participants. Some youth attend 1 day per week, while others go to the program every day after school (i.e., 5 days per week).

Duration. Youth also vary in their duration of attendance. Duration summarizes the history of attendance. Intensity and duration are distinct, as they focus on different time frames. Intensity addresses current attendance, whereas duration addresses the history of attendance in years or terms of a program. For example, of two children currently attending a program 3 days per week, one child may have attended for 3 years, while the other is attending for the first time. In this instance, the two children have exactly the same attendance intensity but differ in terms of duration.

Breadth. Children have many competing responsibilities and opportunities during their out-of-school time. Many children, in fact, attend multiple programs or activities during the week or within the school year. Yet most studies on OST activities focus on youth’s participation in only one activity. Participation in other programs or informal endeavors is typically ignored. Even when researchers use an experimental design and randomly assign youth to either a program or control group, the control group’s attendance in other after school activities is often not discussed. Breadth of attendance refers to the variety of activities that youth attend within and across programs. Some youth obtain breadth by attending multiple OST activities throughout the week, while others experience breadth within their regular OST programs.

Many programs, like Los Angeles’ Better Educated Students for Tomorrow (LA’s BEST), incorporate breadth by offering children a variety of activities (e.g., reading time, sports, dance) within their 5-day-a-week programs. Other programs specialize in one activity, such as baseball or dance. In this case, children can achieve breadth by participating in more than one program. Some project-based learning opportunities within OST programs enable participants to experience a breadth of activities within a single project.
Measuring the Relationship Between Attendance and Outcomes

The 27 studies included in this brief represent programs that offer a wide range of activities, assess attendance in multiple ways, and examine an array of youth outcomes associated with attendance. This section details the specific ways that intensity, duration, and breadth have been measured and how attendance is associated with youth outcomes.

Intensity

As noted above, intensity has been measured in a variety of ways, including the number of days, number of hours, and percentage of available program days that youth attend. The table below lists the ways programs and researchers have used various types of intensity measures to create groupings of youth in their assessment of outcomes associated with attendance.

Overall, attendance intensity has generally been found to be positively associated with many academic and non-academic outcomes, including the following:

- Higher academic achievement and grades
- Spending more time on homework
- Long-term educational and occupational outcomes, such as higher occupational expectations and university enrollment
- Beliefs concerning school, such as higher belief that cheating is bad, and a feeling of belonging at school
- Lower problem behavior
- Less cigarette and drug use
- Higher beliefs about abilities
- Engagement in more community service or volunteering
- Better emotional adjustment, increased happiness, and lower suicidal risk
- More optimistic perceptions of the future

### INTENSITY MEASURES USED IN STUDIES OF OUT-OF-SCHOOL TIME PROGRAMS

<table>
<thead>
<tr>
<th>Intensity Measure</th>
<th>Participant Grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hours per Day or per Week</td>
<td>• High Participator = 4 or more hours; Medium Participator = 1-3 hours; Low Participator = 0 hours</td>
</tr>
<tr>
<td>Days or Sessions per Week</td>
<td>• Primary Arrangement is OST Program = 3 or more days; Nonparticipant = less than 3 days</td>
</tr>
<tr>
<td></td>
<td>• Active Participator = 3 or more days; Nonactive Participator = less than 3 days; Nonparticipant = 0 days</td>
</tr>
<tr>
<td></td>
<td>• High, moderate, and low participation based on the distribution of participation within each program</td>
</tr>
<tr>
<td>Days or Sessions in the Last Year</td>
<td>• Participant = 10 or more sessions; Nonparticipant = less than 10 sessions</td>
</tr>
<tr>
<td></td>
<td>• Expert = 11 of 23 sessions; Novice = 0 days</td>
</tr>
<tr>
<td></td>
<td>• High Participator = 35 or more days; Low Participator = less than 35 days</td>
</tr>
<tr>
<td></td>
<td>• Frequent Participators = 104 or more days; Median Participators = 44-103 days (middle school sample)</td>
</tr>
<tr>
<td></td>
<td>• Frequent Participators = 105 or more days; Median Participators = 49-104 days (elementary school sample)</td>
</tr>
<tr>
<td></td>
<td>• High Participator = participated 79% or more of the days; Moderate Participator = less than 79% of the days</td>
</tr>
<tr>
<td></td>
<td>• Highly Active Participator = 80 or more days; Active Participator = 60-79 days; Nonactive Participator = less than 59 days; Nonparticipant = 0 days</td>
</tr>
<tr>
<td>Time on a Scale</td>
<td>• 1 = rarely or never; 2 = less than once a week; 3 = once or twice a week; 4 = every day or almost every day</td>
</tr>
<tr>
<td></td>
<td>• Not at all; a couple of times a year, etc.</td>
</tr>
<tr>
<td></td>
<td>• 1 = 0 hours; 2 = 1-2 hours per week; 3 = 3-5 hours per week; 4 = 5-10 hours per week; 5 = more than 10 hours per week</td>
</tr>
<tr>
<td></td>
<td>• 1 = 0 hours; 2 = 1-5 hours per week; 3 = 6-10 hours per week; 4 = more than 10 hours per week</td>
</tr>
</tbody>
</table>
As the table indicates, researchers and evaluators have developed cutoffs to compare outcomes based on levels of intensity measured in days of attendance per week. This enables programs to understand the benefits of more or less attendance in the short-term. Participants in the Fifth Dimension program, for example, had to attend at least 10 or 11 days to be classified as participants. Thus, students who attended the program less than 10 days were considered as not having attended at all. Posner and Vandell, on the other hand, classified a child’s primary after school arrangement as an OST program if he or she attended at least 3 days per week. Researchers then tested the differences between youth who attended the same activity at least 3 days per week and those who did not. It should be noted that these tests differed from comparisons based on absolute attendance, because researchers used an intensity criterion to exclude youth who only “tested out” a program or attended sporadically. Researchers have found this strategy to be useful in predicting outcomes. For example, Mayer and colleagues found that youth who had attended the Fifth Dimension at least 10 days in the last year had greater gains on a word problem comprehension test than youth who had never attended the program.

Attendance intensity has also been used to define multiple groups of youth, such as high, medium, and low participators. As the table shows, these definitions have led to the development of groupings of youth based on the frequency of attendance as measured in days and sessions. In several cases, research suggests that youth outcomes improve as their level of attendance intensity increases. Comparisons of high participants and nonparticipants have shown that participants who are high attenders have larger gains on math tests and miss less school than nonattenders. In addition, preliminary evidence suggests that moderate and high attenders have better outcomes than low attendees, with lower rates of truancy and drug use, for example.

Other research suggests a further distinction between moderate and high attenders. In many cases, high attenders have better outcomes than youth who attend at low or moderate levels. High attenders have higher school attendance, grades, feelings of enjoyment in school, academic self-esteem, social interactions with peers, problem-solving skills, and likelihood of being arrested than moderate or low attenders. In a handful of studies, such as one done by Pettit and colleagues, researchers found that moderate rather than high amounts of activity-oriented care were optimal. Moderate intensity was related to better academic and nonacademic outcomes than low or high intensity attendance. (See the section on the Curvilinear Model, page 7, for a more detailed description of this pattern of participation effects.) Overall, most researchers found that higher intensity was linked with better outcomes.

As with most of the work on OST activities, attendance intensity is not always related to outcomes. For example, in some of the evaluations, some academic indicators (e.g., achievement, school attendance) did not significantly differ based on the intensity of youth’s attendance. In addition, certain nonacademic outcomes, such as drug use and problem behaviors, were not significantly related to attendance intensity in some studies. Overall, however, most findings suggest that there is a significant relationship between attendance intensity and outcomes.

**Duration**

Attendance duration, or youth’s activity history, has proven to be another fruitful attendance indicator in predicting youth outcomes. As they have with intensity, researchers have used duration in a variety of ways to predict outcomes. Some have used duration to define whether youth have attended an activity at all. Two studies, for instance, required that youth attend an activity for at least 1 or 2 years to be considered participants. Cutoffs by year seem to yield some intriguing results. For example, youth who attended 4-H for at least 1 year were less likely to engage in delinquent behavior, such as using drugs, damaging property, or smoking cigarettes, than youth who attended for shorter periods. Youth development indicators, such as helping others, talking more with adults, having better attitudes toward school, and taking on leadership roles, were also associated with participating for at least 1 or 2 years.

Researchers have also compared the outcomes of youth who attended a program for different durations. Broh, for example, described sports attendance in terms of duration across 2 years, tenth and twelfth grade. Categories included Never Participated, Participated for 1 Year, or Participated for 2 Years. Continued sports attendance in tenth and twelfth grade was related to higher homework completion, school grades, and achievement test scores. Continued sports attendance was also positively related to nonacademic outcomes, such as confidence, feelings about self, and talking with parents and teachers.

Findings from another 4-H evaluation show that certain outcomes were higher if youth attended a program for more than 1 year. 4-H participants, regardless of their...
duration of attendance, were higher than nonparticipants on several academic and youth development outcomes. However, youth who had been at the club for more than a year were better at communicating, more successful at resolving conflict, spent more hours doing homework, had higher grades, and volunteered more at school than youth who had been in the program less than a year. Finally, evaluations of two other programs showed that significant differences in academic achievement did not emerge unless youth had attended for at least 2 years.

A consistent message emerges from the current research on duration: Duration of at least 2 years is positively related to youth outcomes. Larger differences emerge in outcomes as duration increases. However, it is unclear at this point how longer durations, such as 5 years, are associated with indicators of youth development.

Breadth

Of the three attendance indicators, breadth has received the least attention. Few researchers discuss breadth or use it as a predictor of youth outcomes. Baker and Witt studied breadth within a multicomponent program. Specifically, they examined differences in child outcomes based on the number of activities or components in which youth participated within that one program. Findings suggest that elementary school children who participated in three or more different activities had higher grades and academic test scores than nonparticipants or youth who participated in only one or two activities. In fact, the outcomes of youth who participated in one or two activities were not significantly different from nonparticipants’ outcomes.

Combining Intensity, Duration, and Breadth

Up to this point, the relations between youth outcomes and intensity, duration, and breadth have been summarized separately. However, intensity, duration, and breadth each captures a particular aspect of attendance. As such, the three indicators can be combined to characterize and contribute to our overall understanding of participation. Combining indicators may answer questions about the differences in outcomes of youth who have low intensity over long durations versus youth who have high intensity over short durations, for example. Contrasts such as these require researchers to examine combinations of intensity, duration, and/or breadth. To date, only a handful of researchers have combined these indicators into more complex measures of participation.

Evaluators of the After-School Corporation’s After-School Program (TASC) created an indicator that combined duration and intensity. They compared youth based on how long they had attended—0, 1, or 2 years—and on whether their attendance each year was highly active, active, or nonactive. Although reading scores were not associated with TASC attendance, youth who were highly active for 2 years had the highest gains on math test scores and highest increases in school attendance. This group’s gains were followed by gains for youth who were active for 2 years and for youth who were active for 1 year, respectively. Nonactive participants did not achieve significant gains in their math scores.

In a study of the San Francisco Beacons Initiative, evaluators created a number of attendance variables that combined duration and breadth. Researchers measured duration in terms of the number of sessions youth attended (i.e., spring, fall, summer, following the sessions of the school year). They measured breadth in terms of participation in education activities, other activities, or a combination of the two. The various patterns of attendance across these two indicators were related to some interesting patterns in outcomes. Overall, youth who attended the Beacon Centers for three or more sessions but only in education activities only reported increases in school effort as a result of participation.

Other researchers have studied the number of activities in which youth engaged throughout their high school years, essentially merging breadth and duration. Consider two youth, each of whom participated in four activities across high school. Youth A could have participated in four different activities for 1 year, thus attaining high breadth but low duration. Youth B, on the other hand, could have participated in the same sport for 4 years—low breadth but high duration. Although this characterization makes it impossible to untangle breadth and duration, it has lead to some interesting results. The number of activities in which youth participated across high school was positively associated with numerous outcomes, including satisfaction with life, academic achievement, homework completion, youth’s beliefs about their abilities, educational and occupational plans, and university enrollment.

Utility of Intensity, Duration, and Breadth

Are intensity, duration, and breadth more useful than an indicator that distinguishes between youth who do and do not attend? All of these indicators provide unique information about participation, but they are particularly powerful when used in combination. The following examples suggest that using multiple indicators of youth participation may yield more information about the link between attendance and outcomes than
would selecting just one indicator.

In the Baker and Witt evaluation of two OST programs, findings that included breadth of activities yielded a pattern similar to that found when absolute attendance was examined.56 The results including breadth, however, suggest that youth who participated in at least three or four different activities had better outcomes than youth who participated in fewer activities. This information would not have been obtained if breadth had been omitted from the evaluation.

Findings from the 21st Century Community Learning Centers (21st CCLC) evaluation suggest that intensity also had different relationships to outcomes than absolute attendance did.57 While the two attendance indicators were not used to test all outcomes, some interesting data emerged concerning middle school students. With some outcomes, significant differences were found between youth who attended and those who did not attend. Outcomes were similar, however, for moderate and high attenders. For instance, youth participating in the 21st CCLC program received higher ratings on in-class effort from teachers than did nonparticipants. Ratings of effort did not differ significantly, though, depending on whether youth attended at moderate or high amounts. Thus, effort in class was associated with whether youth attended or not, but not with intensity of attendance.

Other outcomes, however, showed no differences based on absolute attendance but showed significant differences based on intensity. Not being picked on by peers, as well as grades in English, for example, were higher for moderate participators than frequent participators. But when all participants were compared to nonparticipants, these significant findings were not present. In terms of class absenteeism, moderate attendance was associated with lower absenteeism, frequent attendance with even lower absenteeism.

How Attendance Relates to Youth Outcomes: Three Models

Three models can help illustrate how attendance, regardless of the dimension being assessed, can relate to youth outcomes.

Threshold Model

The threshold model is the most basic of the three. Essentially, this model, depicted in Figure 1, suggests that youth will benefit if their attendance exceeds a certain level or threshold. In addition, youth who attend either at or above the threshold will have similar outcomes.

In terms of activity attendance, thresholds are set at different places. For example, the studies that describe differences between youth who do and do not attend are testing whether there is a threshold at any attendance, as depicted in Figure 1. But a threshold can be set anywhere. Researchers could set the threshold at participating 75% of the time or at 3 days per week. Figure 1 is a simple threshold model that includes only one threshold. It is possible to have a second, third, or multiple thresholds. The addition of more thresholds would make Figure 1 begin to resemble a staircase.

Several researchers have tested the threshold model by dividing youth into groups based on attendance. The research on attendance intensity suggests that there may be thresholds at any attendance, at moderate attendance, and at high attendance. For instance, high participators (i.e., those with high attendance) have better academic and social outcomes than moderate participators.58

For two reasons it is difficult at this point to state whether there is a threshold for attendance intensity. First, as demonstrated in the table, these studies differ greatly in how they define the threshold for low, moderate, and high attendance. Thus, even if we found a threshold at “high” attendance, that will mean different amounts of time, depending on the researchers’ characterizations. Second, many of these studies did not test the linear relationship between attendance and outcomes. It is therefore unclear whether outcomes are respectively better for high, moderate, and low participators because there are thresholds or because there is a positive linear relationship between intensity and outcomes. Studies that include tests of the threshold and linear models would help clarify this issue.

A noteworthy number of studies on duration suggest that there may be a threshold at 2 years and possibly at 1 year. Many of the researchers showed that youth outcomes increased after 1 or 2 years of attendance. As with the studies on intensity, it is unclear at this point if there truly is a threshold or if these findings represent the beginning of a linear relationship. It could be that if these relationships were examined over longer durations, we might find that duration and youth outcomes are linearly related or possibly curvilinearly related.

FIGURE 1: THRESHOLD MODEL
Linear Model
A second possible model proposes that the relationship between attendance and outcomes is linear. As Figure 2 depicts, the linear model suggests that as attendance increases, outcomes will increase. According to this model, the more time youth spend in O Ste activities, the better the outcomes. In the strict sense this model suggests that the benefits of attendance do not level off. Youth outcomes keep getting better as attendance increases. For instance, participating 10 hours per week in a program would be associated with better outcomes than participating 8 hours per week.

Most of the research reviewed, particularly concerning attendance intensity, tested the linear model. There are a number of studies suggesting that as intensity increases, so do outcomes. Attendance intensity is associated with academic achievement, improvement of problem behaviors, and emotional adjustment. Most of the research reviewed, particularly concerning attendance intensity, tested the linear model. There are a number of studies suggesting that as intensity increases, so do outcomes. Attendance intensity is associated with academic achievement, improvement of problem behaviors, and emotional adjustment.

Curvilinear Model
A variation on the linear model is the curvilinear model. In the linear model, higher attendance should always lead to better outcomes. The curvilinear model depicted in Figure 3, however, suggests that moderate attendance is associated with good outcomes, while too little or too much attendance is disadvantageous. Too little attendance may not be enough to impact youth outcomes; on the other hand, too much attendance in an activity may be disadvantageous, because other beneficial pursuits and opportunities may be neglected.

A handful of studies present evidence that suggests the association between attendance intensity and outcomes may be curvilinear. In the traditional curvilinear model (Figure 3), moderate intensity is good, but too little or too much is bad. Thus, intensity that is too low has negative relations to outcomes similar to the relations to outcomes of intensity that is too high. The research on O Ste activities, however, suggests that the curvilinear relations between attendance intensity and outcomes may not look like the perfect inverted U-shaped curve presented in Figure 3. Rather, the relations may look more like the curve presented in Figure 4.

In this figure the relations between outcomes and attendance intensity still peak at moderate amounts. The difference between the models is that outcomes are higher for high attenders than low attenders. Many of the curvilinear findings suggest that outcomes are lower for youth participating at high versus moderate intensity but higher for those participating at high versus low intensity.

Curvilinear relations have been found between intensity and several outcomes, including educational/occupational aspirations, substance abuse, test scores, university enrollment, and relationships with peers. Although there is some evidence for a curvilinear relationship between intensity and outcomes, there is little or no evidence for curvilinear relationships between outcomes and either duration or breadth. However, few researchers have rigorously tested the curvilinear model.
Collecting Meaningful Participation Data

We started by asking how much participation, in what kinds of programs, and for which participants is necessary to improve outcomes for youth. As this review has highlighted, the lack of common measurement of participation and how it relates to outcomes renders impossible general statements about how much is enough across all OST programs. Further, since OST programs operate for varying lengths of time each day, varying numbers of days per week, and varying numbers of weeks per year, there is no single “dosage” that will provide optimal impact; we also know that some outcomes require more dosage or intensity to observe impacts. However, while current research does not point to a particular amount of intensity, duration, or breadth that will create beneficial youth outcomes, it is clear from the studies reviewed that meaningful participation in OST programs has beneficial effects.

Simply examining differences between participants and nonparticipants glosses over many of the important aspects of participation. Given the heightened importance of attendance as an indicator of program success, understanding nuanced differences in levels of participation can help programs build data-driven arguments of program effectiveness. Additionally, understanding attendance patterns provides insights to program leaders using evaluation information for program quality improvement. While it is unlikely that the field will develop a single dosage measure that works for all programs, programs can and should collect meaningful attendance data to feed into a system of accountability and program improvement.

But just what participation information should researchers and evaluators collect? In the best of all worlds, they would collect information on all indicators of attendance, as well as information on involvement and engagement. However, the reality of resource and time constraints makes this scenario unlikely. Therefore, program leaders need to work with their evaluation teams to consider seriously which indicators are feasible to collect and which will have the largest benefits to the program. As Fiester asserts, “the ‘right’ methods for collecting, organizing, and analyzing data depend on how program leaders expect to use it—what questions they need to answer and for whom—as well as the program’s size, structure, and resources.” A key component of determining the “right” methods is considering the utility of collecting information on intensity, duration, and breadth.

When discussing the relations between participation in OST activities and outcomes, the issues of youth engagement and involvement must also be addressed. Many programs have struggled with recruiting and retaining participants, and engaging them in meaningful participation. Lauver, Little, and Weiss reviewed promising strategies to attract and sustain participation in OST programs and propose 10 promising strategies, including recruiting participants’ friends, employing energetic and enthusiastic staff, and implementing enticing activities.

Understanding attendance in OST programs is a necessary precursor to understanding how participation affects youth outcomes. This brief is part of a series of HFRP publications designed to support a theoretical model to explain participation. If you would like to be updated on our participation work, consider subscribing to our out-of-school time updates email at www.gse.harvard.edu/hfrp/subscribe.html.

Acknowledgements

Preparation of this brief was made possible through the support of the Charles Stewart Mott Foundation and the W. K. Kellogg Foundation. Special thanks to Jean B. Grossman, Elizabeth Reisner, and Christopher Wimer, who offered a number of insightful comments to improve the paper. We also wish to thank the eight organizations that comprise the Nellie Mae Education Foundation’s Out of School Matters! regional cluster for reading the paper and providing feedback at a May 2004 cluster meeting. Their practitioner insights were invaluable in our review process.
Notes


3 Little, P. M. D., & Harris, E. (2003).


9 Anderson-Butcher et al., 2003.


26 Marsh & Kleitman, 2002.
27 Anderson-Butcher et al., 2003.
34 Schustack et al., 1997.
37 Anderson-Butcher, 2002.
38 Anderson-Butcher et al., 2003.
40 White et al., 2001.
41 Pettit et al., 1997.
42 Anderson-Butcher, 2002.
43 Anderson-Butcher et al., 2003.
44 M arsh & Kleitman, 2002.
47 White et al., 2001.
48 Anderson-Butcher et al., 2003.
49 Marsh & Kleitman, 2002.
51 Anderson-Butcher et al., 2003.
53 Anderson-Butcher et al., 2003.
54 Marsh & Kleitman, 2002.
56 Anderson-Butcher et al., 2003.
59 Welsh et al., 2002.
60 Anderson-Butcher et al., 2003.
61 Marsh & Kleitman, 2002.
62 Anderson-Butcher et al., 2003.
63 Marsh & Kleitman, 2002.
64 Baker & Witt, 1996.
68 White et al., 2001.
69 Pettit et al., 1997.
70 Anderson-Butcher, 2002.
71 Anderson-Butcher et al., 2003.
72 M arsh & Kleitman, 2002.
74 Marsh & Kleitman, 2002.
75 Anderson-Butcher et al., 2003.
76 Marsh & Kleitman, 2002.
78 Anderson-Butcher et al., 2003.
79 Marsh & Kleitman, 2002.
81 Marsh & Kleitman, 2002.
82 Anderson-Butcher, 2002.
83 M arsh & Kleitman, 2002.
84 Marsh & Kleitman, 2002.
85 Anderson-Butcher et al., 2003.
88 Anderson-Butcher, 2002.
89 Marsh & Kleitman, 2002.
90 Marsh & Kleitman, 2002.
91 Anderson-Butcher et al., 2003.
92 Marsh & Kleitman, 2002.
94 Anderson-Butcher et al., 2003.
95 Marsh & Kleitman, 2002.
96 Anderson-Butcher et al., 2003.
97 Marsh & Kleitman, 2002.
98 Anderson-Butcher et al., 2003.
100 Anderson-Butcher et al., 2003.
102 Anderson-Butcher et al., 2003.
103 Marsh & Kleitman, 2002.
104 Anderson-Butcher et al., 2003.
105 Marsh & Kleitman, 2002.
106 Anderson-Butcher et al., 2003.
107 Marsh & Kleitman, 2002.
110 Anderson-Butcher et al., 2003.
111 Marsh & Kleitman, 2002.
112 Anderson-Butcher et al., 2003.
113 Marsh & Kleitman, 2002.
114 Anderson-Butcher et al., 2003.
115 Marsh & Kleitman, 2002.
116 Anderson-Butcher et al., 2003.
117 Marsh & Kleitman, 2002.
118 Anderson-Butcher et al., 2003.
119 Marsh & Kleitman, 2002.
120 Anderson-Butcher et al., 2003.
121 Marsh & Kleitman, 2002.
122 Anderson-Butcher et al., 2003.
123 Marsh & Kleitman, 2002.
124 Anderson-Butcher et al., 2003.
125 Marsh & Kleitman, 2002.
126 Anderson-Butcher et al., 2003.
127 Marsh & Kleitman, 2002.
128 Anderson-Butcher et al., 2003.
129 Marsh & Kleitman, 2002.
130 Anderson-Butcher et al., 2003.
131 Marsh & Kleitman, 2002.
132 Anderson-Butcher et al., 2003.
133 Marsh & Kleitman, 2002.
134 Anderson-Butcher et al., 2003.
135 Marsh & Kleitman, 2002.
136 Anderson-Butcher et al., 2003.
137 Marsh & Kleitman, 2002.
138 Anderson-Butcher et al., 2003.
139 Marsh & Kleitman, 2002.
140 Anderson-Butcher et al., 2003.
141 Marsh & Kleitman, 2002.
142 Anderson-Butcher et al., 2003.
143 Marsh & Kleitman, 2002.
144 Anderson-Butcher et al., 2003.
146 Anderson-Butcher et al., 2003.
147 Marsh & Kleitman, 2002.
148 Anderson-Butcher et al., 2003.
149 Marsh & Kleitman, 2002.
150 Anderson-Butcher et al., 2003.
151 Marsh & Kleitman, 2002.
152 Anderson-Butcher et al., 2003.
154 Anderson-Butcher et al., 2003.
155 Marsh & Kleitman, 2002.
156 Anderson-Butcher et al., 2003.
157 Marsh & Kleitman, 2002.
158 Anderson-Butcher et al., 2003.
159 Marsh & Kleitman, 2002.
160 Anderson-Butcher et al., 2003.
161 Marsh & Kleitman, 2002.
162 Anderson-Butcher et al., 2003.
163 Marsh & Kleitman, 2002.
164 Anderson-Butcher et al., 2003.
165 Marsh & Kleitman, 2002.
166 Anderson-Butcher et al., 2003.
168 Anderson-Butcher et al., 2003.
169 Marsh & Kleitman, 2002.
170 Anderson-Butcher et al., 2003.
171 Marsh & Kleitman, 2002.
172 Anderson-Butcher et al., 2003.
174 Anderson-Butcher et al., 2003.
175 Marsh & Kleitman, 2002.
176 Anderson-Butcher et al., 2003.
177 Marsh & Kleitman, 2002.
178 Anderson-Butcher et al., 2003.
179 Marsh & Kleitman, 2002.
180 Anderson-Butcher et al., 2003.
181 Marsh & Kleitman, 2002.
182 Anderson-Butcher et al., 2003.
183 Marsh & Kleitman, 2002.
184 Anderson-Butcher et al., 2003.
185 Marsh & Kleitman, 2002.
186 Anderson-Butcher et al., 2003.
Research papers and program evaluations were identified through several searches, including the PsychINFO and ERIC research databases, and the Google search engine on the World Wide Web. We used several keywords, such as after-school activities, out-of-school programs, and extracurricular activities. Numerous published and unpublished papers, including program evaluations, were identified for possible inclusion. Several steps were used to screen the papers for inclusion in this review.

This review focuses on general out-of-school programs and activities. We did not include programs that focus solely on tutoring, mentoring, outward bound/adventure, community prevention, or other intensive/holistic programs (e.g., a program that includes social work services). We included out-of-school activities that took place during the school year. We did not include programs that occurred largely during school hours or only in the summer, nor did we include conference presentations, doctoral dissertations, or master's theses. We used only studies with quantitative results that included tests of statistical significance.

To remain focused on current activities and programs, we included studies that have been published or completed since 1990. The number of studies dating before 1990, additionally, was small. Eighty-three studies met all of the above criteria. From this group, 27 studies included intensity, duration, and/or breadth in their analyses. For a list of these studies, see Appendix B.

APPENDIX A: A NOTE ON METHODOLOGY

We included studies that incorporated one of the following methodological designs:

- Experimental. Children are randomly assigned to the OST program or to a control group (i.e., a group consisting of children who did not attend the program).

- Matched comparison group. A study in which youth are not randomly assigned to the OST program; those youth in the program, however, are compared to nonprogram children with whom they were “matched” based on various demographics and/or outcomes.

- Unmatched comparison group. These were similar to the matched comparison group design in that the youth were not randomly assigned to the program. In this design, however, researchers did not try to match program and nonprogram youth.

- Pre-post design. Data was collected on the program youth before and after the program.

APPENDIX B: STUDIES THAT MET THE METHODOLOGICAL AND DESIGN CRITERIA


Gilman, R. (2001). The relationship between life satisfaction, social interest, and frequency of extracurricular activities


---

ABOUT HARVARD FAMILY RESEARCH PROJECT

Founded in 1983 by Dr. Heather Weiss, Harvard Family Research Project conducts research about programs and policies that serve children and families throughout the United States. Publishing and disseminating its research widely, HFRP plays a vital role in examining and encouraging programs and policies that enable families and communities to help children reach their potential.