June 2009

# WHAT WORKS?  TARGETED TRUANCY AND DROPOUT PROGRAMS IN MIDDLE AND HIGH SCHOOL

Children who do not attend school on a regular basis represent a diverse group whose behavioral, mental health, family, and economic problems vary greatly (Hammond, Linton, Smink, & Drew, 2007). Regardless of the background and reasons, however, consistently missing school is a sign of disengagement that is a risk factor for further deterioration in school performance (Epstein & McPartland, 1976).  In fact, truancy has been reliably linked to dropping out (Kaplan, Peck, & Kaplan, 1995; Alexander, Entwisle, & Horsey, 1997) and its consequences (e.g., greater unemployment rates, criminality, etc.; Robins & Ratcliff, 1980; Hibbett, Fogelman, & Manor, 1990; Rouse, 2007).  Interventions targeting truant students seek to avoid these negative outcomes by re-engaging students in academic settings.

In recent years, policymakers have expressed interest in adopting "evidence-based" intervention programs.  With respect to truancy, the goal is to improve the systems that serve truant youth by implementing programs and policies that have been shown to work, as well as eliminate programs that have failed to produce the desired outcomes.  For this purpose, the Institute was asked to evaluate evidence-based intervention and prevention programs for truancy.  Because truancy and school dropout are closely linked, we were instructed to examine dropout prevention programs as well.[1]

This research is part of a larger study of truancy, which examines school- and court-based interventions in Washington and their costs.  In other reports, we review topics that may be important considerations in implementing a specific intervention approach (e.g., cost); however, in this report, we focus on the evidence for effectiveness as a key consideration.

**Summary**

In 2008, the Washington State Institute for Public Policy was directed by the Legislature to study various aspects of truancy.  In the following report, we focus on findings regarding evidence-based practices for truancy reduction and dropout prevention among middle and high school students.  Programs implemented by schools, courts, and law enforcement agencies were considered.

Based on a national review of the literature, we conclude that:

- There are few rigorous studies evaluating the effects of targeted truancy and dropout programs on at-risk students.  In this analysis, only 22 (out of 200) studies met our criteria for rigor.

- Overall, targeted programs for older student populations make small positive impacts on (1) dropping out, (2) achievement, and (3) presence at school (attendance/enrollment).

- When programs are divided based on their central focus or modality, alternative educational programs (e.g., schools-within-schools) and mentoring programs are found to be effective.

- Specifically, Career Academies—an alternative program model that offers a strong career and technical focus—positively impact all three outcomes, as well as high school graduation.

- Alternative schools—separate buildings with specialized academic and other services for at-risk students—have a small *negative* effect on dropping out: more at-risk students drop out of alternative schools than other educational programs.  Additional research is required to better understand this finding.

- Only one rigorous court-based program evaluation was located; thus, this analysis cannot inform court policy or practices.  Because of the key role of the juvenile courts in addressing truancy in many states, additional well-designed studies are imperative.

---

[1] These instructions were provided in consultation with legislative staff and are based on the legislative assignment in ESHB 2687, § 610 (19), chapter 329, Laws of 2008.

## METHODS

Below, we briefly explain the premise behind meta-analysis and its utility for evaluating multiple studies of program evaluations. Afterwards, we outline the criteria used in selecting studies for this analysis.

### What is Meta-Analysis?

Researchers have developed a set of statistical tools to facilitate systematic reviews of the evidence. The set of procedures is called "meta-analysis," which we employed to evaluate the literature on truancy reduction and dropout prevention programs. Importantly, we considered all studies—published and unpublished—that could be located, and did not "cherry pick" studies in advance. Evaluations that met specific criteria were subject to formal statistical testing procedures to determine whether the weight of the evidence indicated significant effects and in which direction (i.e., positive or negative).

### Criteria for Inclusion in the Review

The research deemed appropriate for this review met three broad criteria. For more specific information on inclusion and exclusion criteria (as well as coding rules and statistical formulas for the analyses), see Appendix A.

1) ***Program Focus & Setting.*** Because this analysis is part of a larger investigation of Washington's truancy laws, we focused on truancy and dropout programs that fit within the requirements of the laws and could be implemented by at least one of the systems involved (i.e., schools, courts, or law enforcement). For instance, we reviewed targeted school-based interventions and alternative educational programs, which often serve students at risk of dropping out (Aron, 2006). Court and police programs for truant youth were also considered.

   Programs not included in this analysis are:

   - Programs carried out by institutions not at the heart of truancy laws, such as social service, mental health, or nonprofit agencies in the community[2] (however, if a program demonstrated close collaboration with the schools or courts, it was included);

- Early childhood and elementary school programs[3] (because, similar to national patterns, truancy in Washington becomes especially prevalent in later grades (Office of Superintendent of Public Instruction, 2009) and dropping out is legally permitted only in high school;[4]

- Programs for populations that were identified as at-risk for negative school outcomes due to general factors, such as minority or socioeconomic status;[5]

- School reform or restructuring, including programs for modification of school climate (such as bullying or violence reduction programs);[6]

- "Character programs" or social-emotional learning programs, which target civic contribution, social and conflict resolution skills, and emotional and behavioral regulation first and foremost;[7]

- Programs with a focus after high school graduation (e.g., college enrollment programs for low-income students)[8] or after dropping out (e.g., dropout retrieval or recovery programs);[9] and

- Delinquency or behavior improvement programs for youth who exhibit disruptive behavior and engage in criminal activities (regardless of whether they have been adjudicated).[10]

The exclusion of the programs above is not a judgment regarding their efficacy or relevance. We are aware that many of the groups targeted by these programs overlap with the truant population (for instance, truancy is often seen as the gateway to delinquency, and many children who commit delinquent acts attend school

---

[2] E.g., Boys and Girls Clubs of America; Quantum Opportunity Program (QOP); Adopt-A-Student; Learning, Earning, and Parenting (LEAP); and Cognitive-Behavioral Treatment (CBT) for School Refusal.

[3] E.g., High/Scope Perry Preschool Project, Seattle Social Development Project (SSDP), and Chicago Child-Parent Center Program.
[4] RCW 28A.225.010.
[5] E.g., School Transitional Environmental Program (STEP), Readiness to Learn, and Summer Motivational Academic Residual Training (SMART).
[6] E.g., Comer School Development Program, Talent Development High Schools, First Things First, and Olweus Bullying Prevention Program; also excluded were residential educational settings that generally serve students for reasons other than truancy or even poor achievement.
[7] E.g., Why-Try, Ripple Effects Whole Spectrum Intervention System, Wyman Teen Outreach Program, and Navigation 101.
[8] E.g., Advancement Via Individual Determination (AVID), Upward Bound, and Gaining Early Awareness and Readiness for Undergraduate Programs (GEAR UP).
[9] E.g., National Guard Youth ChalleNGe Corps.
[10] Positive Action Through Holistic Education (PATHE), Preparation Through Responsive Educational Programs (PREP), and CASASTART.

inconsistently; Herrenkohl, Hawkins, Chung, Hill, & Battin-Pearson, 2001; Loeber & Farrington, 1999). Thus, it is possible that investment in programs outside the scope of the current investigation is wise for Washington State; this question is best addressed with additional research. The following analysis seeks to inform only decisions about programs and policies squarely within the provisions of Washington's truancy laws.

2) *Evaluation Design & Methodology.* Not all research is of equal methodological quality, which influences the confidence that can be placed in interpreting the results of a program. Some studies are well-designed and implemented; in such cases, we are more confident that any outcomes were caused by the intervention. However, in less rigorous studies, we have less confidence in the causal effect of the program.

To be included in this analysis, evaluations must have compared outcomes for students who participated in a program with those of an equivalent group that did not participate in the program. The groups did not necessarily have to result from random assignment, but the evaluation must have shown that any comparison group is indeed similar to the treatment group on pre-existing key variables (such as being overage for one's grade, attendance patterns, and academic achievement) that could influence outcome measures. If there were small pre-existing group differences or differences on other variables (e.g., demographics), the authors must have controlled for these differences in their analysis. When such controls are employed, we have more confidence that the outcomes are truly due to the intervention.

A study may utilize a comparison group design, but encounter problems with high attrition (see Gaps in Research sidebar). Studies that suffered from excessive or biased attrition were excluded from this analysis because the original group makeup was judged to be compromised. We do not consider studies that follow the changes of a single treatment group over time to be reliable.

3) *Outcomes.* Finally, the evaluations included in this review had to employ quantifiable measures assessing at least one of the following outcomes: school attendance, high school graduation,[11] or dropout status. Where available, enrollment status was also recorded and combined with attendance into a single variable reflecting students' *presence at school*. In addition, where available, we recorded academic achievement outcomes in the form of standardized test scores and grades.[12]

## FINDINGS

Below is a description of the literature search process and the final set of studies included in the analysis. The program diversity reflected by these studies is outlined. Then, the overall effects of targeted truancy and dropout programs are reviewed. Finally, the effects are divided by classes of programs in order to begin to ascertain which types of interventions are more effective than others for at-risk student populations.

### Studies Included in the Meta-Analysis

Overall, 877 programs were reviewed through websites, databases, and over 460 individual publications. The majority of programs were implemented outside of Washington State; however, two programs were local.[13] Of the 877 programs, 341 met the criterion for appropriate program focus and setting. Two hundred evaluations assessing the effects of these 341 programs were located; however, only 22 studies ultimately met the criteria for research methodology and relevant outcomes. These 22 studies form the basis of our analysis.

Among the final 22 studies (for citations, see Appendix B), many included multiple outcomes and some used multiple samples. For instance, one study investigated 13 different programs in various locations throughout the country. We coded each sample and outcome separately,

---

[11] This outcome includes only students who earned high school diplomas and does not take into account students who earned GED's.
[12] Studies in which outcomes were measured using self-report—for instance, student participants reporting on their own attendance— were excluded. Previous studies have shown that these measures are unreliable for at-risk student populations (e.g., Gleason & Dynarski, 1994, p. 81).
[13] Middle College High School (Dynarski, Gleason, Rangarajan, & Wood,1998); Summer Training and Education Program (Grossman & Sipe, 1992).

such that ultimately there were 100 effects in the analysis. These effects are divided by outcome in the following way:

- Presence at school (enrollment/ attendance) = 35

- Dropout = 30

- Achievement (grades/test scores) = 29

- High school graduation = 6

Notice that valid graduation outcomes[14] were obtained in only six sites (three publications). Graduation outcomes are difficult to collect because they require careful documentation of highly mobile populations and long-term follow-up by researchers; both problems are explained in detail below (see Gaps in Research sidebar).

In all, the studies represent 34 distinct programs, which are diverse in the types of services provided. For examples of these services, see Exhibit 1.

More than one service was offered in most programs (e.g., a program could offer contingency management and parent outreach). Also, programs delivered the services in various settings (e.g., alternative school, school-within-school, after school, etc.). There were several common combinations of services and settings, which formed the basis for larger program classes, described below. Of note, out of 34 programs, only one court intervention was represented.

*Exhibit 1*
**Services Provided by Truancy and Dropout Programs in the Analysis**

| Services | Description |
| --- | --- |
| Academic remediation/ tutoring | Assess students' academic skill deficiencies and provide specialized or intensive instruction to improve competency. |
| Career/ technical education | Increase student awareness of the connection between school and work life, and teach technical skills (e.g., through vocational courses, employer internships, etc.). |
| Case management | Problem-solve barriers to school success and refer student and family to community or other services, based on needs. |
| Contingency management | Systematically reward desirable behaviors (e.g., on-time attendance) and punish undesirable behaviors (e.g., not completing homework). |
| Counseling | Analyze and problem-solve barriers to school success, including personal, family, and social challenges, in a safe, supportive environment. |
| Mentoring/ advocacy | Provide students with a role model who supports their educational endeavors and advocates for their success in the school system. |
| Monitoring attendance | Intensively track student attendance and follow up with student and parents to prevent tardies/absences. |
| Parent outreach | Engage parents in identifying and solving their child's school problems; often families are referred to social or other supportive services. |
| Youth development | Provide opportunity for skill-building, horizon expansion, competence, and resilience, and improve connections to school and positive adults. |
| Additional services | Offer services that meet additional needs of the at-risk population served (e.g., childcare center/parenting classes, school-based health center). |

[14] In some studies, graduation is reported, but we judged that it was not a valid outcome (e.g., because data on graduation were missing for a significant portion of the original program participants) and, therefore, did not include it in the analysis.

## EVIDENCE-BASED PROGRAMS

When results from all of the truancy- and dropout-specific programs in our analysis are combined, we find modest but positive statistically signficant impact on dropping out, achievement, and presence at school.[15] Graduation effects are moderate;[16] however, there are only six effect sizes representing three types of programs. Thus, at this time, it is not clear that all (or even most) truancy and dropout programs have the same effects on graduation. Because results reflect the effects of programs that differ greatly in their approach, setting, and intensity, we further investigated which interventions are and are not effective.

Due to the limited number of rigorous studies, evidence for specific programs was not available (with the exception of Career Academies, which boast a relatively high number of research replications). Thus, we grouped programs by their general focus or modality. For example, in all mentoring programs, students are paired with an adult who is expected to support the child's educational endeavors, advocate for their success in school, and connect them to appropriate services.

For each program class, the results reflect the evidence-based effect expected for an "average" implementation of such a program. Of course, we recognize that each class of programs is diverse and some programs may ultimately prove to be more effective than others.

---

[15] Adjusted average effect sizes for each outcome are as follows: dropping out = .054 (p=.07); achievement = .048 (p=.01); and presence at school = .098 (p=.00).
[16] The average effect size = .158 (p=.00).

In Exhibit 2, the plus and minus signs indicate a statistically significant effect on the indicated school outcome. Plus signs designate a positive effect, such as greater achievement or less dropping out, whereas negative effects, such as less presence at school, are shown with a minus sign. Zeroes indicate that the outcome was measured with at least two samples; however, the program effect is not statistically significant (i.e., there is no reliable effect).[17]

*Exhibit 2*
**Effects of Truancy and Dropout Programs for Middle and High School Students on School Outcomes**

| Program Class | Dropout | Presence at School (Enrollment & Attendance) | Achievement (Test Scores & Grades) | Graduation |
|---|---|---|---|---|
| **Alternative educational programs:** Programs involving a group of students *in a traditional school* (e.g., school-within-a-school) that usually offer small class size, more individualized instruction, and/or different instructional methods and material (e.g., vocational curriculum). | **+** | **+** | **+** | **+** |
| **Mentoring:** Providing students with positive role models, who help with specific academic issues (e.g., homework), advocate for the student in the school system, and connect them to other services (e.g., social services). | **+** | **+** | **0** | **0** |
| **Behavioral programs:** Targeting students' school behaviors by helping them analyze and problem-solve negative behaviors, and/or by establishing a system of contingencies (rewards, punishments) for desirable and undesirable behaviors. | N/A[*] | **+** | **0** | N/A[*] |
| **Youth development:** Preventing negative school outcomes by promoting bonding with positive figures and school environment, fostering competence and skill building, and supporting resilience. Rather than focusing on remediation of youth's weaknesses, programs target healthy development and build on youth's strengths.^ | N/A[*] | **0** | **0** | N/A[*] |
| **Academic remediation:** Providing students with additional or intensive instruction to improve academic skills, usually in core subject areas (e.g., reading, math). | **0** | **0** | **0** | N/A[*] |
| **Alternative schools:** Schools with *separate facilities and services* for students who struggle in traditional school settings. Schools usually incorporate an alternative curriculum (often academic remediation) and psychosocial services (e.g., counseling, case management). | **–** | **0** | **0** | **0** |

**\*** This table summarizes the results of multiple evaluations per program class. Results are not shown when fewer than two studies or two outcomes were identified for a particular program class. Complete results are provided in Appendix C.

^Roth & Brooks-Gunn, 2003.

---

[17] When a program has a statistically significant effect, we are at least 90 percent confident that the result is not due to chance alone.

Three intervention classes improve student attendance and enrollment (i.e., presence at school):

- Alternative educational programs,
- Behavioral programs, and
- School-based mentoring.

Alternative educational programs (housed within traditional schools) and mentoring programs also significantly reduce students' dropping out from school. Achievement—measured by grades and test scores—and high school graduation were only improved by alternative educational programs.

While the effects were statistically significant, they were modest in size. For instance, one of the larger program effects observed was .203, the effect size for student dropout rates in mentoring programs. In this high-risk group of students, 35 percent of the comparison group dropped out. With this effect size, 28 percent of students in the mentoring group would be expected to drop out, an overall reduction of 7 percent.

Notably, we found that at-risk students attending alternative schools, which have separate facilities, are significantly **more** likely to drop out than similarly at-risk students in traditional schools. The magnitude of this effect is 4 percent, with 31 percent of traditional school students and 35 percent of alternative school students dropping out.

No positive outcomes were observed for alternative schools, academic remediation, or youth development programs. For a complete presentation of effect sizes and other statistical results, see Appendix C.

### Alternative educational programs

Only alternative programs had a positive effect on all four outcomes. These programs provide specialized instruction to a group of students *within* a traditional school, often separating them for at least some of their academic courses and integrating them with other students for elective classes. In this sense, alternative programs differ from alternative schools, in which the entire school day is spent in separate facilities that often include different rules and norms from traditional schools. In some cases, alternative programs espouse a unique focus (for instance, career training) or instructional method (computer-assisted learning).

In this meta-analysis, the positive effects of alternative programs can be attributed to a particular intervention model known as Career Academies,[18] in which small learning communities are formed within a larger high school. They combine an academic and technical curriculum around a career theme (which differs based on local interest), and establish partnerships with community-based employers to provide work-based learning opportunities. A unique feature of Career Academies is that they serve not only struggling students, but also seek to include achieving students (Kemple & Rock, 1996). According to the Career Academy Support Network (CASN), there are 6,000 to 8,000 Academies in the US,[19] with 14 currently known to be in operation in Washington (J. Patrick, personal communication, June 11, 2009; L. Holland, personal communication, June 16, 2009).

### Mentoring programs

Mentoring programs pair struggling students with an adult who serves as a role model, supports school achievement, and helps the youth navigate an often complex school system. This study found that such programs make a small positive impact on school presence and dropping out, but not on achievement.[20] Importantly, most of the interventions evaluated here employed paid mentors. Such compensation may have incentivized the mentors to perform better than volunteer mentors, who are more typically utilized in the community.

---

[18] Weighted average effect sizes for Career Academies were all statistically significant: dropout = .147 (p=.01); presence at school = .113 (p=.00); achievement = .060 (p=.02); and graduation = .248 (p=.00). In contrast, the average effect sizes for all other alternative programs were not statistically significant (dropout = .053; presence in school = .110, achievement = .118), indicating no reliable effect of the latter on outcomes.

[19] An exact figure does not exist because there is no federal oversight or documentation of Career Academies. The total estimate is based on a 2004 national survey, figures from established networks of Academies, and CASN knowledge of trends on the ground (C. Dayton, personal communication, June 11, 2009).

[20] There were not enough evaluations measuring the effect of mentoring programs on graduation to draw conclusions.

### *Negative findings for alternative schools*

As explained earlier, these specialized schools offer separate facilities intended to serve students whose needs are not met by traditional schools. Although in theory such schools could serve highly gifted youth or children with specialized interests (e.g., art), it has been found that the majority of alternative schools in the United States (Kleiner, Porch, & Farris, 2002), including Washington (Baker, Gratama, & Bachtler, 2007), serve students at risk for school failure. Thus, many offer remedial instruction, mental health services, case management, and specialized on-site services (e.g., childcare for offspring).

Exhibit 2 shows that alternative schools do not have an impact on school presence, achievement, or graduation rates. More importantly, the data reveal that alternative schools have small negative effects on dropping out, that is, slightly more students in alternative schools drop out (35 percent) than do otherwise similar students in other settings (traditional school, alternative programs; 31 percent).

This finding may sound counterintuitive, as alternative schools are designed to address the unique academic and psychosocial needs of students who do not succeed in traditional schools; however, the interventions provided may not be responsible for increased dropout rates. For instance, alternative schools may inadvertently lead to deviant peer influence, whereby—despite adult supervision—youth learn antisocial behavior from one another in a group setting that isolates them from the influence of prosocial adolescents (Dishion, McCord, & Poulin, 1999; Reinke & Walker, 2006). Another possibility is that alternative schools, especially those accepting students with behavioral problems, are not schools of choice. It has been suggested that when students perceive transfer to an alternative school as punishment, rather than a choice to attend a more appropriate educational environment, their academic motivation may be hurt (Raywid, 1994). Lack of motivation may, in turn, impact the decision to remain in school.[21]

Washington State has a large number of alternative schools relative to other states (National Center for Education Statistics, 2008); thus the finding regarding their effect on dropout rates may be of interest to educators and policymakers. Note, however, that given the small number of studies, we cannot comment on any particular structure, philosophy, or intervention feature of alternative schools. Given the great variability in the ways that alternative schools select students and carry out programs, it is possible that some alternative schools in Washington curb dropout rates while others have little impact or amplify the rates.

---

[21] Friedrich (1997) found that alternative schools/programs with choice have lower rates of dropout than traditional programs, but schools/programs without choice have higher dropout rates than their traditional counterparts. These results are based on few studies and analyses did not control for methodological factors or school characteristics; therefore, definitive conclusions cannot be drawn at this time.

## CONCLUSIONS

This investigation centered on interventions targeting middle and high school students who are chronically truant or at-risk for dropping out of school. Furthermore, in this study, service providers were limited to systems involved in Washington's truancy laws: schools, courts, and law enforcement. Twenty-two publications representing a wide range of interventions were included.

Importantly, in this literature, it is clear there are very few rigorous studies of interventions to prevent truancy and dropout. Creative interventions have been attempted by talented individuals on the ground; however, without research evaluations, little can be made of their effectiveness. Moreover, many evaluations that have been conducted have employed poor research designs such that one can have little confidence that the intervention caused the measured outcomes.

The studies included in this analysis, judged to be sufficiently rigorous, showed that targeted programs have small positive effects on dropping out, presence at school, and student achievement. Specifically, the evidence points to alternative programs, mentoring approaches, and to a lesser extent, behavioral interventions as those that hold promise for at-risk populations.

It is informative that alternative programs, housed within traditional schools, were effective, but alternative schools, in separate facilities, did not improve student outcomes. This finding may reflect a need to maintain some level of integration among at-risk and high-achieving students.

Although we found that Career Academies constituted an effective model for alternative programs, in other areas, we were unable to identify brand-name programs or intervention features that were especially successful. Further research on specific programs is necessary because the interventions contained within each program class are diverse, and it is possible that some programs are more effective than others. Additional research could reveal, for instance, which mentoring programs are most effective in re-engaging students, improving their grades, and ensuring their graduation from high school.

Additional research could also shed light on the finding that alternative schools lead to slightly higher rates of dropping out. By highlighting different school philosophies—program implementation, student selection, etc.—such evaluations would reveal which features are responsible for the negative effect. (At the same time, this research might isolate features that lead to greater student success in alternative schools.)

Finally, research is lacking on court-based interventions. In this review of the literature, we identified only one court program that was rigorously evaluated. Because Washington's laws clearly imbue the juvenile courts with a role in reducing truancy, additional evidence on the effectiveness of court policies and practices is important.

To this end, the Institute will be conducting an evaluation of the truancy petition process statewide. We will examine the academic and criminal outcomes of students who undergo this process in comparison to students with similar attendance records who do not have a court petition. A report with the findings will be published in July 2009.

It is important to stress that while this meta-analysis was comprehensive in locating published and unpublished studies of targeted interventions for truant and at-risk youth in older age groups, its scope was limited. Programs not examined in this analysis due to a different focus, setting, or age group may also be effective in reducing truancy and dropping out (e.g., delinquency interventions, community-based approaches, early childhood education). The legislature may wish to further investigate these possibilities as it seeks to improve the outcomes of Washington's youngest residents.

## Appendix A
## Meta-Analytic Procedures

To identify evidence-based programs (EBP) in reducing truancy and preventing dropout, we conducted reviews of the relevant research literature. In recent years, researchers have developed a set of statistical tools to facilitate systematic reviews of evaluation evidence. The set of procedures is called "meta-analysis"; we employ this methodology in our study.[22] In Appendix A, we describe these general procedures, as well as the unique coding for research rigor carried out in order to determine whether further analyses (e.g., multiple regressions) were appropriate.

### A1. Study Selection and Coding Criteria
A meta-analysis is only as good as the selection and coding criteria used to conduct the study.[23] Following are the key choices we made and implemented.

**Study Selection.** We searched for all truancy reduction and dropout prevention program evaluation studies conducted since 1970. The studies had to be written in English. We used three primary means to identify and locate these studies: (a) we consulted the programs and corresponding evaluations listed by centers or organizations that have an expertise in truancy and/or dropout, for example, the National Dropout Prevention Center/Network (NDPC/N), National Center for School Engagement (NCSE), American Youth Policy Forum (AYPF), and the What Works Clearinghouse of the U.S. Department of Education Institute of Education Sciences (IES); (b) we read systematic and narrative reviews of the truancy, dropout, and related research literatures (such as school refusal due to anxiety disorders); (c) we examined the citations in the individual studies; and (d) we conducted independent literature searches of research databases using search engines such as PsycInfo, ERIC, and Google. As we describe, the most important inclusion criteria in our study was that an evaluation have a control or comparison group. Therefore, after first identifying all possible studies using these search methods, we attempted to determine whether the study was an outcome evaluation that had a comparison group. If a study met these criteria, we then secured a paper or electronic copy of the study for our review.

**Peer-Reviewed and Other Studies.** We examined all program evaluation studies we could locate with these search procedures. Some of these studies were published in peer-reviewed academic journals, while many others were found in reports obtained from government, private, or nonprofit agencies. It is important to include non-peer reviewed studies, because it has been suggested that peer-reviewed publications may be biased to show positive program effects. Therefore, our meta-analysis includes all available studies we could locate regardless of published source.

**Control and Comparison Group Studies.** Our analysis only includes studies that had a control or comparison group. That is, we did not include studies with a single-group, pre-post research design. This choice was made because it is only through rigorous comparison group studies that average

treatment effects can be reliably estimated. We do include quasi-experimental observational studies that are of sufficient statistical rigor.

**Random Assignment and Quasi-Experiments.** Random assignment studies were preferred for inclusion in our review, but we also included non-randomly assigned control groups. We only included quasi-experimental studies if sufficient information was provided to demonstrate comparability between the treatment and comparison groups on important pre-existing conditions such as school attendance, achievement, and grade retention (essentially, indicators of participants' school engagement and performance).

**Enough Information to Calculate an Effect Size.** Following the statistical procedures in Lipsey and Wilson (2001), a study had to provide the necessary information to calculate an effect size. If the necessary information was not provided, the study was not included in our review.

**Mean-Difference Effect Sizes.** For this study, we coded mean-difference effect sizes following the procedures in Lipsey and Wilson (2001). For dichotomous measures, we used the D-cox transformation to approximate the mean difference effect size, as described in Sánchez-Meca, Marín-Martínez, and Chacón-Moscoso (2003). We chose to use the mean-difference effect size rather than the odds ratio effect size because we frequently coded both dichotomous and continuous outcomes (odds ratio effect sizes could also have been used with appropriate transformations).

**Unit of Analysis.** Our unit of analysis for this study was an independent test of a treatment for a particular sample. Some studies reported results for multiple samples, which were often located at different sites. We included each sample as an independent observation if a unique and independent comparison group was used for each.

**Multivariate Results Preferred.** Some studies presented two types of analyses: raw outcomes that were not adjusted for covariates such as age, gender, or other pre-intervention characteristics, and those that had been adjusted with multivariate statistical methods. In these situations, we coded the multivariate outcomes.

**Outcomes of Interest.** Because this is a study of truancy reduction and dropout prevention programs, evaluations could only be included if they contained at least one of these two outcomes (i.e., attendance or dropping out/graduation rates). In addition, enrollment status and measures of achievement, such as test scores and grades, were included in the analysis in order to better understand how programs impact several educational outcomes. Ultimately, based on associations among outcomes and numbers of studies included per outcome, two sets of outcomes were combined: attendance and enrollment became "presence at school," and "test scores and grades" became "achievement."

**Averaging Effect Sizes.** When a study reported multiple test scores, attendance measures, or subject grades, we calculated the effect size for each measure and then computed a weighted average effect size (e.g., the effect on average test scores). In combining outcomes into broader categories (i.e., presence at school and achievement), if a

---

[22] We follow the meta-analytic methods described in Lipsey and Wilson (2001).

[23] All studies used in the meta-analysis are identified in the references in Appendix B of this report. Many other studies were reviewed but did not meet the standards set for this analysis.

study reported both of the component outcomes (e.g., both test scores and grades), we used the weighted average of the two effect sizes. In studies that reported only one of the component outcomes, it became the final outcome (e.g., test scores became achievement).

**Longest Follow-Up Periods.** When a study presented outcomes with varying follow-up periods, we generally coded the effect size for the longest follow-up period. This allows us to gain the most insight into the long-run impact of various treatments. Occasionally, we did not use the longest follow-up period if it was clear that outcomes for the longer reported follow-up period were unreliable, for instance, because a dramatically increased attrition rate adversely affected the treatment and comparison group samples.

**Special Coding Rule for Effect Sizes: Declaration of Significance by Category.** Most studies in our review had sufficient information to code exact mean-difference effect sizes. Some studies, however, reported some, but not all the information required. We followed the following rule for these situations: Some studies reported results of statistical significance tests in terms of categories of p-values. Examples include: p<=.01, p<=.05, or non-significant at the p=.05 level. We calculated effect sizes for these categories by using the most conservative p-value in the category. Thus, if a study reported significance at p<=.05, we calculated the effect size at p=.05. This is the most conservative strategy. If the study simply stated a result was non-significant, we computed the effect size assuming a p-value of .50.

## A2. Procedures for Calculating Effect Sizes

Effect sizes measure the degree to which a program has been shown to change an outcome for program participants relative to a comparison group. There are several methods used by meta-analysts to calculate effect sizes, as described in Lipsey and Wilson (2001). In this analysis, we used statistical procedures to calculate the *mean difference effect sizes* of programs. We use the standardized mean difference effect size for continuous measures and the D-cox transformation as described in Sánchez-Meca et al. (2003, equation 18) to approximate the mean difference effect size for dichotomous outcome variables.

A(1): $d_{Cox} = \ln\left(\dfrac{p_e(1-p_c)}{p_c(1-p_e)}\right)/1.65$

In equation (A1), $d_{cox}$ is the estimated effect size, which is derived by dividing the log odds ratio by the constant 1.65. $Pe$, represents the percentage outcome for the experimental or treatment group and, $Pc$, is the percentage outcome for the control group.

For continuous outcome measures, we use the standardized mean difference effect size statistic (Lipsey & Wilson, table B.10, equation 1).

A(2): $ES_m = \dfrac{M_e - M_c}{\sqrt{\dfrac{SD_e^2 + SD_c^2}{2}}}$

In this formula, $ES_m$ is the estimated effect size for the difference between means from the research information; $M_e$ is the mean number of an outcome for the experimental group; $M_c$ is the mean number of an outcome for the control group; $SD_e$ is the standard deviation of the mean number for the experimental group; and $SD_c$ is the standard deviation of the mean number for the control group.

Often, research studies report the mean values needed to compute $ES_m$ in (A2), but they fail to report the standard deviations. Sometimes, however, the research will report information about statistical tests or confidence intervals that can then allow the pooled standard deviation to be estimated. These procedures are also described in Lipsey and Wilson.

**Adjusting Effect Sizes for Small Sample Sizes.** Since some studies have very small sample sizes, we follow the recommendation of many analysts and adjust for this. Small sample sizes have been shown to upwardly bias effect sizes, especially when samples are less than 20. Following Hedges (1981), Lipsey and Wilson (2001, equation 3.22) report the "Hedges correction factor," which we use to adjust all mean difference effect sizes ($N$ is the total sample size of the combined treatment and comparison groups):

A(3): $ES'_m = \left[1 - \dfrac{3}{4N-9}\right] \times \left[ES_m, or, ES_{m(p)}\right]$

**Computing Weighted Average Effect Sizes, Confidence Intervals, and Homogeneity Tests.** Once effect sizes are calculated for each program effect, the individual measures are summed to produce a weighted average effect size for a program area. We calculate the inverse variance weight for each program effect and these weights are used to compute the average. These calculations involve three steps. First, the standard error, $SE_m$ of each mean effect size is computed with (Lipsey & Wilson, 2001, equation 3.23):

A(4): $SE_m = \sqrt{\dfrac{n_e+n_c}{n_e n_c} + \dfrac{(ES'_m)^2}{2(n_e+n_c)}}$

In equation (A4), $n_e$ and $n_c$ are the number of participants in the experimental and control groups and $ES'_m$ is from equation (A3).

For dichotomous outcomes, the standard error, $SEd_{cox}$, is computed with (Sánchez-Meca et al., 2003, Equation 19):

A(5): $SE_{d_{Cox}} = \sqrt{0.367\left[\dfrac{1}{O_{1E}} + \dfrac{1}{O_{2E}} + \dfrac{1}{O_{1C}} + \dfrac{1}{O_{2C}}\right]}$

In equation (A5), $O_{1E}$ and $O_{1C}$ represent the success frequencies of the experimental and control groups. $O_{2E}$ and $O_{2C}$ represent the failure frequencies of the experimental and control groups.

Next, the inverse variance weight $w_m$ is computed for each mean effect size with (Lipsey & Wilson, equation 3.24):

A(5): $\quad w_m = \dfrac{1}{SE_m^2}$

The weighted mean effect size for a group of studies in program area $i$ is then computed with (Lipsey & Wilson, p. 114):

A(6): $\quad \overline{ES} = \dfrac{\sum (w_{m_i} ES'_{m_i})}{\sum w_{m_i}}$

Confidence intervals around this mean are then computed by first calculating the standard error of the mean with (Lipsey & Wilson):

A(7): $\quad SE_{\overline{ES}} = \sqrt{\dfrac{1}{\sum w_{m_i}}}$

Next, the lower, $ES_L$, and upper limits, $ES_U$, of the confidence interval are computed with (Lipsey & Wilson):

A(8): $\quad \overline{ES_L} = \overline{ES} - z_{(1-\alpha)}(SE_{\overline{ES}})$

A(9): $\quad \overline{ES_U} = \overline{ES} + z_{(1-\alpha)}(SE_{\overline{ES}})$

In equations (A8) and (A9), $z_{(1-\alpha)}$ is the critical value for the $z$-distribution (1.96 for $\alpha = .05$).

The test for homogeneity, which provides a measure of the dispersion of the effect sizes around their mean, is given by (Lipsey & Wilson, p. 116):

A(10): $\quad Q_i = (\sum w_i ES_i^2) - \dfrac{(\sum w_i ES_i)^2}{\sum w_i}$

The Q-test is distributed as a chi-square with $k$-1 degrees of freedom (where $k$ is the number of effect sizes).

**Computing Random Effects Weighted Average Effect Sizes and Confidence Intervals.** When the p-value on the Q-test indicates significance at values of p less than or equal to .05, a random effects model is performed to calculate the weighted average effect size. This is accomplished by first calculating the random effects variance component, $v$ (Lipsey & Wilson, 2001, p. 134).

A(11): $\quad v = \dfrac{Q_i - (k-1)}{\sum w_i - (\sum wsq_i / \sum w_i)}$

This random variance factor is then added to the variance of each effect size and finally all inverse variance weights are recomputed, as are the other meta-analytic test statistics.

## A3. Institute Adjustments to Effect Sizes for Methodological Quality, Outcome Assessment Period, and Researcher Involvement

In Appendix C, we show the results of the meta-analysis, calculated with the standard meta-analytic formulas described in Appendix A2. In the last column, we list the "Adjusted Effect Size." These adjusted effect sizes, which are derived from the unadjusted results, are always smaller than or equal to the unadjusted effect sizes we report in the same exhibit.

In Appendix A3, we describe our rationale for making these downward adjustments. In particular, we make three types of adjustments that are necessary to better estimate the results that we are more likely to achieve in real-world settings. We make adjustments for: (a) the methodological quality of each study; (b) the assessment period of the outcomes; and (c) the degree to which the researcher(s) who conducted each study were invested in the program's design.

**Research Design Quality.** Not all research is of equal quality, and this greatly influences the confidence that can be placed in the results of a study. Some studies are well designed and implemented, and the results can be viewed as accurate representations of whether the program itself worked. Other studies are not designed as well, and less confidence can be placed in any reported differences. In particular, studies of inferior research design cannot completely control for sample selection bias or other unobserved threats to the validity of reported research results. This does not mean that results from these studies are of no value, but it does mean that less confidence can be placed in any cause-and-effect conclusions drawn from the results.

To account for the differences in the quality of research designs, we use a 5-point scale as a way to adjust the reported results. The scale is based closely on the 5-point scale developed by researchers at the University of Maryland (Sherman, Gottfredson, MacKenzie, Eck, Reuter, & Bushway, 1998, chap. 2). On this 5-point scale, a rating of "5" reflects an evaluation in which the most confidence can be placed. As the evaluation ranking gets lower, less confidence can be placed in any reported differences (or lack of differences) between the program and comparison or control groups.

On the 5-point scale as interpreted by the Institute, each study is rated with the following numerical ratings:

- A **"5"** is assigned to an evaluation with well-implemented random assignment of subjects to a treatment group and a control group that does not receive the treatment/program. A good random assignment study should also indicate how well the random assignment actually occurred by reporting values for pre-existing characteristics for the treatment and control groups.

- A **"4"** is assigned to a study that employs a rigorous quasi-experimental research design with a program and matched comparison group, controlling with statistical methods for self-selection bias that might otherwise influence outcomes. These quasi-experimental methods may include estimates made with a convincing instrumental-variables modeling approach, or a Heckman approach to modeling self-selection (Rhodes, Pelissier, Gaes, Saylor, Camp, & Wallace, 2001). A level 4 study may also be used to

"downgrade" an experimental random assignment design that had problems in implementation, perhaps with significant attrition rates.

- A **"3"** indicates a non-experimental evaluation where the program and comparison groups were reasonably well matched on pre-existing differences in key variables. There must be evidence presented in the evaluation that indicates few, if any, significant differences were observed in these salient pre-existing variables. Alternatively, if an evaluation employs sound multivariate statistical techniques (e.g., logistic regression) to control for pre-existing differences, and if the analysis is successfully completed, then a study with some differences in pre-existing variables can qualify as a level 3.

- A **"2"** involves a study with a program and matched comparison group where the two groups lack comparability on pre-existing variables and no attempt was made to control for these differences in the study.

- A **"1"** involves a study where no comparison group is utilized. Instead, the relationship between a program and an outcome, i.e., drug use, is analyzed before and after the program.

We do not use the results from program evaluations rated as a "1" on this scale, because they do not include a comparison group and, thus, no context to judge program effectiveness. We also regard evaluations with a rating of "2" as highly problematic and, as a result, do not consider their findings in the calculations of effect. In this study, we only considered evaluations that rated at least a 3 on this 5-point scale.

An explicit adjustment factor is assigned to the results of individual effect sizes based on the Institute's judgment concerning research design quality. This adjustment is critical and the only practical way to combine the results of a high-quality study (a level 5) with those of lesser design quality (level 4 and level 3 studies). The specific adjustments made for these studies are based on our knowledge of research in education and other topic areas that show that random assignment studies (i.e., level 5 studies) have, on average, smaller absolute effect sizes than weaker-designed studies (Lipsey, 2003; Cox, Davidson, & Bynum, 1995). Thus, we use the following "default" adjustments to account for studies of different research design quality:

- A level 5 study carries a factor of 1.0 (that is, there is no discounting of the study's evaluation outcomes).

- A level 4 study carries a factor of .75 (effect sizes discounted by 25 percent).

- A level 3 study carries a factor of .50 (effect sizes discounted by 50 percent).

- We do not include level 2 and level 1 studies in our analyses.

These factors are subjective to a degree; they are based on the Institute's general impressions of the confidence that can be placed in the predictive power of evaluations of different quality.

The effect of the adjustment is to multiply the effect size for any study, $ES'_m$, in equation (A3) by the appropriate research design factor. For example, if a study has an effect size of -.20, and it is deemed a level 4 study, then the -.20 effect size would be multiplied by .75 to produce a -.15 adjusted effect size.

**Adjusting Effect Sizes of Studies With Short-Term Follow-Up Periods.** Short-term outcomes are potentially problematic because they may not reflect the full effect of the program. For example, dropout outcomes, measured immediately following a program ending in 9th grade, may not capture whether the student ultimately withdrew from high school. To reflect the lower level of confidence we can ascribe to measures of dropout in the earlier grades, a discount factor of .75 was applied (i.e., the effect size was reduced by 25 percent). On the other hand, dropout outcomes measured during 11th or 12th grade were presumed to have assessed the majority of research participants in the intervention and comparison group who experienced these outcomes (i.e., those who will have dropped out).

Presence at school and achievement outcomes were not discounted because, unlike dropping out, it was less clear that there was an endpoint at which assessment was optimal (such as 12th grade).

For the fourth outcome, high school graduation, we did not include studies that measured graduation if all of the sample could not have been reasonably expected to graduate. For example if graduation was measured for a group of students in 10th, 11th and 12th grade, we would not expect those in 10th and 11th grade to have yet graduated.

**Generalizability: Researcher Involvement in Program Design and Implementation.** The purpose of the Institute's work is to identify and evaluate programs that can make beneficial improvements to Washington's actual service delivery system. There is some evidence that programs closely controlled by researchers or program developers have better results than those that operate in "real world" administrative structures.[24] This difference may be due to the evaluator/developer's inordinate investment in the program or personal characteristics that are not easily replicated (e.g., charisma).

Based on this evidence, we distinguish between evaluations in which the roles of researcher and developer or implementer overlap (and therefore there is concern regarding generalizability) and those in which different individuals are clearly responsible for these roles. As a parameter for all studies deemed not to be "real world" trials, the Institute discounts $ES'_m$ by .5, although this can be modified on a study-by-study basis.

---

[24] Lipsey (2003) found that, for juvenile delinquency evaluations, programs in routine practice (i.e., "real world" programs) produced effect sizes only 61 percent as large as research/demonstration projects. For inability to replicate positive results of a highly controlled efficacy study of a school-based substance abuse program in "real world" settings, see Hallfors, Cho, and Sanchez (2006). In child/adolescent mental health, see Weisz, Huey, and Weersing (1998).

## Appendix B
## Citations of Studies Used in the Meta-Analysis of Truancy and Dropout Programs

Brooks, B. D. (1975). Contingency management as a means of reducing school truancy. *Education, 9*, 206-211.

Bry, B. H. & George, F. E. (1979). Evaluating and improving prevention programs: A strategy from drug abuse. *Evaluation and Program Planning, 2*, 127-136.

Bry, B. H. & George, F. E. (1980). The preventive effects of early intervention on the attendance and grades of urban adolescents. *Professional Psychology, 11*(2), 252-260.

Cox, S. M. (1999). An assessment of an alternative education program for at-risk delinquent youth. *Journal of Research in Crime and Delinquency, 36*(3), 300-322.

DeSocio, J., VanCura, M., Nelson, L. A., Hewitt, G., Kitzman, H., & Cole, R. (2007). Engaging truant adolescents: Results from a multifaceted intervention pilot. *Preventing School Failure, 51*(3), 3-9.

*Dynarski, M. & Wood, R. (1997). *Helping high-risk youth: Results from the Alternative Schools Demonstration Program*. Princeton, NJ: Mathematica Policy Research.

*Dynarski, M., Gleason, P., Rangarajan, A., & Wood, R. (1998). *Impacts of dropout prevention programs.* Princeton, NJ: Mathematica Policy Research.

Elliott, M. N., Hanser, L. M., & Gilroy, C. L. (2002). Career Academies: Additional evidence of positive student outcomes. *Journal of Education for Students Placed at Risk, 7*(1), 71-90.

Fantuzzo, J., Grim, S., & Hazan, H. (2005). Project START: An evaluation of a community-wide school-based intervention to reduce truancy. *Psychology in the Schools, 42*, 657-667.

Grossman, J. B. & Sipe, C. L. (1992). *Summer Training and Education Program (STEP): Report on long-term impacts.* Philadelphia: Public/Private Ventures.

*Hayward, B. & Tallmadge, G. (1995). *Strategies for keeping kids in school: Evaluation of dropout prevention and reentry projects in vocational education, final report*. Raleigh, NC: Research Triangle Institute.

*Kemple, J. J. & Snipes, J. C. (2000). *Career Academies: Impacts on students' engagement and performance in high school*. New York: Manpower Demonstration Research Corporation.

Langley, K. K. (1993). A program evaluation of Project GRADS: A dropout prevention program. *Dissertation Abstracts International, 54*(05), 1680A. (UMI No. 747490321)

*Larson, K. A. & Rumberger, R. W. (1995). ALAS: Achievement for Latinos through academic success: Project evaluation 1990-1995. In H. Thorton (Ed.), *Staying in school: A technical report of three dropout prevention projects for middle school students with learning and emotional disabilities* (pp. A-1 - A-71). Minneapolis: University of Minnesota, Institute on Community Integration.

*McPartland, J. M. & Nettles, S. M. (1991). Using community adults as advocates or mentors for at-risk middle school students: A two-year evaluation of Project RAISE. *American Journal of Education, 99*(4), 568-586.

Miller, D. (1986). Effect of a program of therapeutic discipline on the attitude, attendance, and insight of truant adolescents. *Journal of Experimental Education, 55*(1), 49-53.

Rhodes, H. J. (2007). *Confronting the challenges of student engagement: A case study of a school-based intervention*. Santa Monica, CA: RAND.

*Rossi, R. J., DuBois, P. A., McLaughlin, D. H., de Mello, V. B., Goya, K. A., Tallmadge, G. K., et al. (1995). *Evaluation of projects funded by the School Dropout Demonstration Assistance Program: Final evaluation report*. Palo Alto, CA: American Institutes for Research.

Sinclair, M. F., Christenson, S. L., & Thurlow, M. L. (2005). Promoting school completion of urban secondary youth with emotional or behavioral disabilities. *Exceptional Children, 71*(4), 465-482.

Slicker, E. K. & Palmer, D. J. (1993). Mentoring at-risk high school students: Evaluation of a school-based program. *The School Counselor, 40*, 327-334.

Sullivan, P. D. (1982). A comparison of students in an urban mini-school program, designed to retain and improve the reading, English, and mathematics achievement of potential dropouts, with students in a traditional high school program. *Dissertation Abstracts International, 42(7*), 2962A. (UMI No. 748405771)

Webb, C. G. (1973). An evaluation of the Early Evening Vocational-Technical Training Program utilized as an approach to reach the potential dropout. *Dissertation Abstracts International, 34*(06), 3163A. (UMI No. 761279111)

*Based on the three criteria for inclusion (i.e., program focus and setting, evaluation design and methodology, outcomes), not all sites or samples reported in this publication were included in the meta-analysis.

## Appendix C
## Meta-Analytic Estimates of Standardized Mean Difference Effect Sizes:
## Truancy and Dropout Programs for Middle and High School Students

| Type of Intervention Program (and its effect on outcomes included in the meta-analysis) | Number of Effect Sizes Included in Analysis (Number of cases in treatment groups) | Meta-Analytic Results Before Applying Institute Adjustments | | | | | Adjusted Effect Size Used in the Meta-Analysis (estimated effect after adjustments for the methodological quality of the evidence, assessment period, and researcher involvement) | Notes to Table |
|---|---|---|---|---|---|---|---|---|
| | | Fixed Effects Model | | | Random Effects Model | | | |
| | | Weighted Mean Effect Size | | Homogeneity Test | Weighted Mean Effect Size | | | |
| | | ES | p-value | p-value | ES | p-value | ES | |
| **EFFECTS OF ALL PROGRAMS IN THE ANALYSIS:** | | | | | | | | (1) |
| School Dropout | 30 (4014) | 0.055 | 0.108 | 0.001 | 0.096 | 0.071 | **0.054** | |
| Presence at School (Attendance & Enrollment) | 35 (3745) | 0.153 | 0.000 | 0.000 | 0.191 | 0.000 | **0.098** | |
| Achievement (Test Scores & Grades) | 29 (2712) | 0.079 | 0.006 | 0.077 | 0.087 | 0.015 | **0.048** | |
| High School Graduation | 6 (635) | 0.260 | 0.002 | 0.244 | 0.258 | 0.009 | **0.158** | |
| **EFFECTS BY PROGRAM CLASS:** | | | | | | | | (2) |
| **Academic remediation programs:** | | | | | | | | |
| School Dropout | 5 (1365) | 0.050 | 0.418 | 0.005 | 0.267 | 0.139 | **0.000** | |
| Presence at School (Attendance & Enrollment) | 3 (535) | -0.020 | 0.781 | 0.574 | na | na | **0.000** | |
| Achievement (Test Scores & Grades) | 4 (532) | 0.086 | 0.227 | 0.987 | na | na | **0.000** | |
| **Alternative educational programs:** | | | | | | | | |
| School Dropout | 12 (1120) | 0.153 | 0.008 | 0.341 | na | na | **0.093** | |
| Presence at School (Attendance & Enrollment) | 12 (1435) | 0.174 | 0.000 | 0.000 | 0.209 | 0.007 | **0.101** | |
| Achievement (Test Scores & Grades) | 12 (1297) | 0.105 | 0.009 | 0.102 | na | na | **0.050** | |
| High School Graduation | 2 (124) | 0.496 | 0.000 | 0.356 | na | na | **0.248** | |
| **Alternative schools:** | | | | | | | | |
| School Dropout | 7 (883) | -0.118 | 0.098 | 0.158 | na | na | **-0.104** | |
| Presence at School (Attendance & Enrollment) | 5 (704) | 0.101 | 0.151 | 0.470 | na | na | **0.000** | |
| Achievement (Test Scores & Grades) | 3 (205) | 0.003 | 0.977 | 0.638 | na | na | **0.000** | |
| High School Graduation | 3 (440) | 0.161 | 0.167 | 0.867 | na | na | **0.000** | |
| **Behavioral programs:** | | | | | | | | |
| School Dropout | 1 (246) | 0.087 | 0.682 | na | na | na | **0.000** | (3) |
| Presence at School (Attendance & Enrollment) | 6 (375) | 0.191 | 0.011 | 0.000 | 0.512 | 0.020 | **0.223** | |
| Achievement (Test Scores & Grades) | 4 (277) | 0.048 | 0.589 | 0.013 | 0.178 | 0.569 | **0.000** | |
| **Community-based court:** | | | | | | | | |
| Presence at School (Attendance & Enrollment) | 1 (189) | 0.503 | 0.000 | na | na | na | **0.252** | (3) |
| **Mentoring programs:** | | | | | | | | |
| School Dropout | 4 (215) | 0.237 | 0.069 | 0.060 | na | na | **0.203** | |
| Presence at School (Attendance & Enrollment) | 5 (278) | 0.199 | 0.030 | 0.223 | na | na | **0.146** | |
| Achievement (Test Scores & Grades) | 4 (183) | 0.075 | 0.499 | 0.326 | na | na | **0.000** | |
| High School Graduation | 1 (71) | -0.070 | 0.773 | na | na | na | **0.000** | (3) |
| **Vocational programs (not in alternative settings):** | | | | | | | | |
| Presence at School (Attendance & Enrollment) | 1 (15) | 0.916 | 0.011 | na | na | na | **0.458** | (3) |
| **Youth development programs:** | | | | | | | | |
| School Dropout | 1 (185) | -0.327 | 0.199 | na | na | na | **0.000** | (3) |
| Presence at School (Attendance & Enrollment) | 2 (215) | -0.063 | 0.555 | 0.754 | na | na | **0.000** | |
| Achievement (Test Scores & Grades) | 2 (218) | 0.011 | 0.917 | 0.022 | -0.125 | 0.670 | **0.000** | |

(1) Positive effect sizes indicate favorable outcomes (i.e., better school achievement, more presence at school, higher rates of graduation, less dropping out).
(2) All outcomes were not evaluated for each program class; thus, in some classes, outcomes are not reported.
(3) This effect size was not included in Exhibit 2 of the report because findings from one sample are not considered reliable.
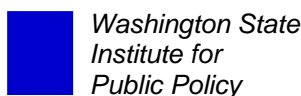
# Bibliography

Alexander, K. L., Entwisle, D. R., & Horsey, C. S. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of Education, 70*(2), 107-127.

Aron, L. Y. (2006). *An overview of alternative education.* Washington, DC: The Urban Institute.

Baker, D. B., Gratama, C. A., & Bachtler, S. D. (2007, July). *Office of Superintendent of Public Instruction Washington State Alternative High School Initiative: A feasibility study.* Bothell, WA: The BERC Group.

Cox, S. M., Davidson, W. S., & Bynum, T. S. (1995). A meta-analytic assessment of delinquency-related outcomes of alternative education programs. *Crime & Delinquency, 41*(2), 219-234.

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54*(9), 755-764.

Dynarski, M., Gleason, P., Rangarajan, A., & Wood, R. (1998). *Impacts of dropout prevention programs.* Princeton, NJ: Mathematica Policy Research.

Ellickson, P. L., Bianca, D., & Schoeff, D. C. (1988). Containing attrition in school-based research. *Evaluation Review, 12*(4), 331-351.

Epstein, J. L., & McPartland, J. M. (1976). The concept and measurement of the quality of school life. *American Educational Research Journal, 13*(1), 15-30.

Friedrich, K. R. (1997). Alternative education for at-risk youth: An analytical review of *evaluation findings* (Doctoral dissertation, Texas A&M University, 1997). *Dissertation Abstracts International, 58*(04), 1197A.

Gleason, P., & Dynarski, M. (1994). *Falling behind: Characteristics of students in federally funded dropout prevention programs. Part one: Targeted projects.* Princeton, NJ: Mathematica Policy Research.

Grossman, J. B., & Sipe, C. L. (1992). *Summer Training and Education Program (STEP): Report on long-term impacts.* Philadelphia: Public/Private Ventures.

Hallfors, D., Cho, H., Sanchez, V. (2006). Efficacy vs effectiveness trial results of an indicated "model" substance abuse program: Implications for public health. *American Journal of Public Health, 96*(12), 2254–2259.

Hammond, C., Linton, D., Smink, J., & Drew, S. (2007). *Dropout risk factors and exemplary programs: A technical report.* Clemson, SC: National Dropout Prevention Center/Network.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*, 107-128.

Herrenkohl, T. I., Hawkins, J. D., Chung, I., Hill, K. G., & Battin-Pearson, S. (2001). School and community risk factors and interventions. In R. Loeber & D. P. Farrington (Eds.), *Child Delinquents* (pp. 211-246). Thousand Oaks, CA: Sage.

Hibbett, A., Fogelman, K., & Manor, O. (1990). Occupational outcomes of truancy. *British Journal of Educational Psychology, 60*(1), 23-36.

Institute of Education Sciences. (2008). *IES Practice Guide: Dropout Prevention.* Washington, DC: Author.

Kaplan, D. S., Peck, B. M., & Kaplan, H. B. (1995). A structural model of dropout behavior: A longitudinal analysis. *Applied Behavior Science Review, 3*(2), 177-193.

Kemple, J. J. & Rock, J. L. (1996). *Career Academies: Early implementation lessons from a 10-site evaluation.* New York: Manpower Demonstration Research Corporation.

Kleiner, B., Porch, R., & Farris, E. (2002, September). *Public alternative schools and programs for students at risk of education failure: 2000-01.* Washington, DC: National Center for Education Statistics.

Lehr, C. A., Hansen, A., Sinclair, M. F., & Christenson, S. L. (2003). Moving beyond dropout towards school completion: An integrative review of data-based interventions. *School Psychology Review, 32*(3), 342-364.

Lipsey, M. W. (2003). Those confounded moderators in meta-analysis: Good, bad, and ugly. *The Annals of the American Academy of Political and Social Science, 587*(1), 69-81.

Lipsey, M. W., & Wilson, D. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Loeber, R., & Farrington, D. P. (1999). *Serious and Violent Juvenile Offenders.* Thousand Oaks, CA: Sage.

National Center for Education Statistics, Digest of Education Statistics (2008). *Table 98: Public elementary and secondary schools, by type and state or jurisdiction: 1990-91, 2000-01, and 2006-07.* Retrieved June 11, 2009, from http://nces.ed.gov/programs/digest/d08/tables/dt08_098.asp

Office of Superintendent of Public Instruction. (2009). *Reports to the legislature, truancy Becca Bill, 2007-08.* Retrieved from http://www.k12.wa.us/Truancy/default.aspx

Raywid, M. A. (1994). Alternative schools: The state of the art. *Educational Leadership, 52*(1), 26-31.

Reinke, W. M., & Walker, H. M. (2006). Deviant peer effects in education. In K. A. Dodge, T. J. Dishion, & J. E. Lansford (Eds.), *Deviant peer influences in programs for youth: Problems and solutions* (pp. 122-140). New York: The Guilford Press.

Rhodes, W., Pelissier, B., Gaes, G., Saylor, W., Camp, S., & Wallace, S. (2001). Alternative solutions to the problem of selection bias in an analysis of federal residential drug treatment programs. *Evaluation Review, 25*(3), 331-369.

Robins, L. N., & Ratcliff, K. S. (1980). The long-term outcome of truancy. In I. Berg & L. Hersov (Eds.), *Out of school* (pp. 85-110). Chichester, UK: John Wiley & Sons.

Roth, J. L., & Brooks-Gunn, J. (2003). What exactly is a youth development program? Answers from research and practice. *Applied Developmental Science, 7*(4), 94-111.

Rouse, C. E. (2007). Consequences for the labor market. In C. R. Belfield & H. M. Levin (Eds.), *The price we pay: Economic and social consequences of  inadequate education* (pp. 99-124). Washington, DC: Brookings Institution Press.

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods, 8*(4), 448–467.

Shannon, G. S., & Bylsma, P. (2005).  *Promising programs and practices for dropout prevention.* Olympia, WA: Office of Superintendent of Public Instruction, Assessment and Research.

Sherman, L., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1998). *Preventing crime: What works, what doesn't, what's promising.* College Park: University of Maryland, Department of Criminology and Criminal Justice.

Sinclair, M. F., Christenson, S. L., & Thurlow, M. L. (2005). Promoting school completion of urban secondary youth with emotional or behavioral disabilities. *Exceptional Children, 71*(4), 465-482.

Tyler, J. H. (2008). *Brief Number 2: Dropout prevention programs: What research has to say about what works* (Virginia Family Impact Seminar #1). Retrieved April 1, 2009, from http://www.familyimpactseminars.org/s_vafis01c02.pdf

Weisz, J. R., Huey, S. J., & Weersing, V. R. (1998). Psychotherapy with children and adolescents: The state of the art. In T. H. Ollendick & R. J. Prinz (Eds.), *Advances in clinical child psychology, Volume 20* (pp. 49–91). New York: Plenum Press.