A meta-analysis of the effect of juvenile delinquency interventions on academic outcomes ☆

Janay B. Sander *, Erika A. Patall, Laura A. Amoscato, Alexandra L. Fisher, Catherine Funk

The University of Texas at Austin, United States

ABSTRACT

This meta-analysis examined the effects of juvenile delinquency interventions on academic outcomes. After retrieving over 250 reports, 15 reports met inclusion criteria and provided 134 effect sizes (92 unadjusted and 42 adjusted) based on 20 separate samples in a variety of settings, including school, community, and juvenile justice settings. Heterogeneity of the samples, generally weak research designs, and the absence of control conditions in many recovered reports was a limitation in the existing research. Overall, there were limited positive effects of juvenile delinquency interventions on academic outcomes. The lack of theory-driven or empirically supported academic interventions was notable. Studies with the weakest designs produced the largest effects on academic achievement, and school attendance outcomes were enhanced only for older adolescents. The implications of findings for future research and policy are discussed.

Published by Elsevier Ltd.

1. Introduction

The concept of the *school to prison pipeline* is one of the most pressing concerns related to juvenile delinquency and education. This term refers to the phenomenon in which students gradually become disengaged from school while simultaneously becoming more involved in crime and delinquency (Christle, Jolivette, & Nelson, 2005). The overall cost to society is a concern. That is, youths (5–17 years) account for 18.9% of the total United States population (Howden & Meyer, 2011), but are responsible for approximately 16% of all violent crime, and 26% of all property crimes (Puzzanchera, 2009). According to the Office of Juvenile Justice and Delinquency Prevention (OJJDP), nearly 2 million youths are arrested each year and the overall incidence and costs that result from acts of delinquency are considerable (Puzzanchera, 2009). Education seems to be a central component of the problem and solution (Council of State Governments and Public Policy Research Initiative, 2011). Low educational attainment is linked to higher rates of delinquency and recidivism (Cottle, Lee, & Heilbrun, 2001; Leone et al., 2003), while increased literacy and educational achievement is associated with lower rates of crime (Cottle et al., 2001; Keith & McCray, 2002).

The educational challenges of populations engaged in juvenile delinquency have been clearly noted in research literature since the 1950s (Wilson, Lipsey, & Soydan, 2003). One often-cited characteristic of juvenile offenders is the high proportion of them who are involved in special education. Special education participation of youths in juvenile justice systems is considerably higher than that of the general population; 35% of juvenile offenders participate in special education (Quinn, Rutherford, Leone, Osher, & Poirier, 2005; Zabel & Nigro 1999) versus 8% national rate of the general population of the United States (National Center for Education Statistics, 2010). However, this characteristic alone is not a useful descriptor. The problem is not that youths are enrolled in special education participation per se; it is low academic skill. The average age of adjudicated youth is 15 years, an age at which most youth are entering the 10th grade, but the average reading level of adjudicated youth is aligned with the 4th grade, where children are typically age 9–10 years (Leone, Krezmien, Mason, & Meisel, 2005; Vacca, 2008). In light of the close associations among reading level, educational attainment, and crime and recidivism rates, academic skill is a highly relevant consideration for juvenile justice rehabilitation efforts.

In spite of the importance of educational attainment, much of the current literature focuses exclusively on behavioral and crime/recidivism outcomes. It is not surprising that there are numerous reports on interventions to reduce crime and improve behavioral outcomes for youth, many of which have examined educational variables and outcomes in some capacity. Yet, despite the clear links between delinquency and education, the literature about juvenile delinquency interventions and educational outcomes is under-developed. Specifically, reports of educational interventions and studies of potential factors that may influence interventions are more or less effective for promoting positive educational outcomes for youth offenders are not clearly summarized in the current literature.

A synthesis of the educational effects of juvenile delinquency interventions would provide a meaningful contribution to inform future interventions and research. To address the gaps in the literature, the primary purpose of this investigation is to synthesize existing research about the effects of juvenile delinquency interventions on academic outcomes. We start by briefly summarizing the research about delinquency interventions not specific to education, summarizing some general patterns that guided our quantitative analyses, and then we describe the available literature about educational effects of delinquency programs.

1.1. Prior meta-analyses on the effects of juvenile delinquency interventions

Several meta-analyses have summarized effects of juvenile delinquency interventions on behavioral outcomes, primarily recidivism and crime. Although educational outcomes have not generally been the focus of these prior meta-analyses (though some have examined academic outcomes to a limited extent or used education variables as moderators), it is useful to have an overview of the characteristics of successful juvenile delinquency interventions that focus on behavioral outcomes. This framework serves as a background to understand what factors might be important to consider in understanding the effects of juvenile delinquency interventions on academic outcomes. One of the most comprehensive datasets of juvenile delinquency reports has yielded multiple meta-analytic investigations (Lipsey, 1999, 2009; Wilson & Lipsey, 2001, 2007). In the most recent update using that database, Lipsey (2009) included 500 published and unpublished original studies conducted between 1950 and 1996. These earlier syntheses of delinquency interventions contribute several helpful points to understand to what extent, when, and how interventions reduce delinquency.

There are four consistent findings related to juvenile delinquency interventions and behavioral outcomes. First, even among delinquency interventions that offer moderate success, recidivism is only reduced by about 12% overall (Lipsey, 1999), with generally small effect sizes ($d = .10$ to $.15$) (Lipsey & Cullen, 2007). Second, the most effective delinquency interventions are those based on theory and a clear approach for change, often with a solid research base. Third, interventions conducted as standard field practice are somewhat helpful, but less so than more intensive implementations of interventions that are closely supervised by researchers. Fourth, positive, as opposed to punitive, behavioral interventions to address aggression or disruptive behaviors typical in juvenile offender populations appear to be helpful, even when delivered as routine programs apart from a rigorous research-driven approach in schools (Lipsey & Cullen, 2007; Wilson, Lipsey, & Derzon, 2003). These four overall findings are clear in terms of behavioral outcomes, and they provide useful guides for additional synthesis for the same population.

In terms of the effects of juvenile delinquency interventions on academic outcomes, it is prudent to consider whether intervention characteristics found to be associated with positive behavioral outcomes may also be associated with positive academic outcomes. That is, would academic outcomes mirror the effect sizes of the behavioral outcomes? Further, we use these prior meta-analyses as a guide in the current meta-analysis because examining similar questions will allow findings from this meta-analysis to be integrated and compared with those of prior syntheses of juvenile delinquency intervention effects.

1.2. Educational programs that address behavior

During the past two decades, there have been a number of studies of school-based interventions geared at improving positive behaviors while also reducing "acting out" and delinquency-type behaviors among the general student population. Many of these studies are showing positive effects on both academic and behavioral outcomes. The effects of those programs have been summarized in two separate, comprehensive synthesis investigations (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Wilson, Gottfredson, & Najaka, 2001), indicating that behavior-focused programs can be successful at addressing both behavioral and academic concerns. One group of programs, the Social and Emotional Learning (SEL) programs, was the target of a recent meta-analysis that included academic and behavioral outcomes in school settings. SEL programs produced moderate positive effects on conduct problems and academic performance ($d = .22$ and $.27$), even when academic intervention was not the primary program component (Durlak et al., 2011). Another approach, Response to Intervention (RTI), a data-driven approach in education to address academic concerns (Edmonds et al., 2009; Vaughn, Denton, & Fletcher, 2010) has emerging support for behavioral concerns as well (Gresham, 2007). Positive Behavior Support programs (PBS; Sugai & Horner 1999) have also been designed to promote academic and behavioral success, often for the same youths that are served by juvenile justice centers.

Overall, research on these school-based universal prevention and targeted early intervention programs indicates that they are helpful in promoting positive academic outcomes, even when the focus is on reducing disruptive and aggressive behavior (Wilson et al., 2001). The body of education research generally shows that school based prevention and early intervention programs, even when the focus is behavioral and not academic, help improve behavior *and* academic achievement. In light of these findings, it would seem important to clarify the effects of interventions with populations engaged in delinquency on academic outcomes because there may be academic benefits even of juvenile delinquency programs that have no academic component.

1.3. Academic intervention research in juvenile corrections

While it would seem that even behavioral and conduct focused interventions can have an impact on academic outcomes, it also seems like common sense to suspect that those juvenile delinquency programs that include an academic component might be most successful at improving academic outcomes. There have been a handful of studies examining the effects of using academic interventions with a reading focus within juvenile correctional detention facilities (e.g. Campbell, Marsh, & Stickel, 1993; Drakeford, 2002; Leone et al., 2005; Malmgren & Leone, 2000). In fact, a research synthesis on the topic was conducted for studies carried out in correctional facility settings (Leone et al., 2005). The scholars could not draw firm conclusions about overall effects of such programs because there were only six studies available which met their inclusion criteria (Krezmien & Mulcahy, 2008; Leone et al., 2005). The authors noted the difficulty of carrying out reading research in detention and corrections settings as one of their main findings.

However, community services and public schools also address that academic needs of juvenile offenders and programs in these settings are relevant to include in a synthesis for this population. Youths involved in juvenile delinquency typically spend less than 30 days in detention facilities, and only about 14% of all youths who are involved in juvenile justice at any given time are in detention. So the great majority of these youth spend the most of their time in community, school, and other service agencies, where they receive intervention (OJJDP, 2008). To date, a number of studies that have examined the effects of academic and reading-specific intervention programs with juvenile offenders outside the detention setting have not been included in a synthesis of research. In addition, academic intervention research for youths in secondary education, the age range typical for youths engaged in juvenile delinquency, has increased within the most recent 15 years (Edmonds et al., 2009). In sum, given the expansion of educational intervention research, as well as the wide range of settings within which intervention for juvenile delinquency is known to occur, there seems to be a pressing need to synthesize and compare

the effects of the most recent juvenile delinquency intervention programs on academic outcomes using a range of approaches and conducted in a range of settings.

Finally, and as previously alluded to, while some past syntheses examining the effects of juvenile delinquency interventions (many of which have already been mentioned) have examined their effects on academic outcomes (Leone et al., 2005; Wilson, Lipsey, & Derzon, 2003; Wilson, Lipsey, & Soydan, 2003), many of the most comprehensive synthesis studies (e.g., Lipsey, 1999, 2009; Lipsey & Cullen, 2007) may not accurately reflect the current state of academic interventions in practice. This would seem to be problematic given that a number of innovations in education research interventions have come about in recent years. To explain more, in the largest set of studies used in meta-analysis of juvenile delinquency interventions (n = 305) that includes academic outcomes, the most recent study was conducted in 1996 (Wilson, Lipsey, & Derzon, 2003; Wilson, Lipsey, & Soydan, 2003). Meanwhile, the most effective reading interventions in education settings, especially with upper elementary and secondary education levels (grades 6–12) are reported later (1990–2010) (e.g. Edmonds et al., 2009; Wanzek, Wexler, Vaughn, & Ciullo, 2010). Notably, studies reporting outcomes of targeted reading interventions in education research have increased dramatically since the mid-1990s (Edmonds et al., 2009; Wanzek, Wexler, Vaughn, & Ciullo, 2010). In sum, recent advances in educational interventions would not yet be captured in syntheses of delinquency intervention literature given the dates of most studies included in previously published meta-analytic studies. As such, it seems important to conduct and up-to-date meta-analysis on the effects of juvenile delinquency interventions on academic outcomes.

1.4. Moderators of interest

Finally, there are several aspects of delinquency interventions that are important to consider as potential moderators of effects. Moderators examined in past syntheses related to juvenile offending or education interventions typically include: research design, demographic information (e.g., age, sex, grade level), the duration of the intervention, the training or type of professional delivering the intervention, and the type of setting (Edmonds et al., 2009; Lipsey, 2009; McCarty & Weisz, 2007; Weisz, McCarty, & Valeri, 2006; Wilson & Lipsey, 2007). All of these factors have been found to explain some of the variability in outcomes for delinquency interventions and/or within the psychosocial and educational intervention literature. As such, we consider these factors in this synthesis of research on juvenile delinquency intervention effects on academic outcomes.

1.4.1. Research design

It is necessary to consider whether the findings are due to the difference between groups or the effects of the intervention, rather than research design characteristics (Weisz et al., 2006; Wilson & Lipsey, 2001). Randomized controlled trials are rare within juvenile delinquency literature, but quasi-experimental designs with appropriate controls are acceptable and are often available in this literature (Lipsey, 2009). Typically, higher quality research designs produce smaller effect sizes and more trustworthy results (Edmonds et al., 2009; Lipsey, 2009; Weisz et al., 2006). We expect that this will be the case in the current synthesis as well. However, the extent to which randomized experiments differ from quasi-experimental designs with or without equating groups remains an important question to address in this synthesis.

1.4.2. Age of participants

In studies about juvenile delinquency, a younger age of arrest is typically predictive of a more problematic level of criminal involvement (Cottle et al., 2001). Older youths tend to respond better to delinquency interventions (Lipsey, 1999). At the same time, however, younger youths respond more rapidly to educational interventions

(Kroesbergen & Van Luit, 2003), while older youths may require more intense and targeted educational interventions when they reach grade 6 or higher (Edmonds et al., 2009). This moderator is of interest, but is considered exploratory in this study.

1.4.3. Gender and ethnicity/race

Gender may also be important, as girls tend to have higher educational achievement and may respond differently to intervention compared to boys in juvenile delinquency target samples (Zabel & Nigro, 2007). Boys are more likely to engage in crime in general (Cottle et al., 2001). In a meta-analytic study conducted by Wilson, Lipsey, and Soydan (2003), delinquency interventions had small differences in effects, but overlapping confidence intervals, for White compared to minority cultural groups. The question of whether there are differences in the effects of intervention between gender and cultural or ethnic groups remains important to consider.

1.4.4. Setting

The type of setting is a moderator of interest that has been included in some prior meta-analyses (Wilson, Lipsey, & Derzon, 2003; Wilson, Lipsey, & Soydan, 2003), but is not always available depending on the focus of the synthesis. In their synthesis of meta-analyses, Lipsey and Cullen (2007) reported that community-based programs have more favorable effects than detention settings, but highlight that the reasons for that could be due to the differences in samples, not settings or programs. This moderator is important in order to understand the effects of various types of services available for rehabilitation and supporting education. In line with prior synthesis in this area of research, we also examine setting as a moderator.

1.4.5. Training and staff characteristics

The type of training, overall qualifications of the staff providing the intervention, supervision and support of staff during implementation, and type of professional delivering the intervention have all been noted moderators (Wilson & Lipsey, 2007) in prior studies that would be relevant to the current study. If an intervention is designed and carried out by a researcher, as opposed to staff who are not directly affiliated with the intervention per se, the effects usually positive and more pronounced (Wilson & Lipsey, 2007). In addition, interventions delivered by staff with higher levels of training or advanced degrees also yield larger effects (Weisz et al., 2006). It is expected that this investigation would identify similar patterns.

1.4.6. Intervention type and dosage

In the intervention literature, including educational, psycho-social or behavioral interventions, the type and duration of intervention also moderate outcomes (McCarty & Weisz, 2007; Wilson & Lipsey, 2007). Specific interventions that target an identified skill deficit or symptom, rather than a more general approach, tend to produce larger effect sizes (Edmonds et al., 2009). For example, a focused reading comprehension intervention rather than a general broad tutoring approach for academic skills, or using cognitive–behavioral therapy as opposed to more general counseling for mental health is recommended. In addition, interventions that are delivered over period of about 6 to 12 months tend to also produce larger effect sizes than briefer interventions (McCarty & Weisz, 2007; Wilson & Lipsey, 2007). The intervention type will be noted with as much specificity as possible, particularly for any academic intervention aspects. The duration will also be noted.

1.5. Summary of current gaps in the literature

The literature examining the effects of juvenile delinquency interventions is currently limited in several ways. Prior meta-analyses have been limited in terms of the setting of studies included in the analyses, and included studies are not reflective of recent advances in

educational interventions. Further, even those meta-analyses that have examined academic outcomes, little attention has been paid to characteristics of the research design, academic intervention, setting, or sample that may influence the effects of juvenile delinquency programs on academic outcomes.

Future research would benefit from a systematic analysis that examines the effect of juvenile delinquency interventions on academic outcomes. In addition, the current study offers more depth in terms of the level of specificity for the academic *and* behavioral/psychosocial intervention components that will be examined. The findings from this study will address which components are most effective for improving academic outcomes of youths engaged in juvenile delinquency. Academic intervention may not be the only program that improves educational attainment, and improvements may occur in school, juvenile justice, or community settings. That is, even program components geared toward behavioral outcomes may improve academic outcomes and may do so across a wide variety of settings.

With these considerations in mind, the goals of this research synthesis were twofold. First, we examined the overall effects of intervention programs designed to reduce juvenile delinquency across school, community, and juvenile justice settings on academic achievement, school attendance, and school attitudes. Prior syntheses of research that examined academics as a specific outcome of juvenile delinquency intervention effects were either including studies conducted in juvenile detention facilities only, or in schools only, but not across the multiple settings where juvenile delinquency programs are offered. Second, in this meta-analysis, we investigated several moderators that could potentially influence the effectiveness of delinquency interventions on academic outcomes. We predicted that juvenile delinquency programs that include more rigorous designs, more highly trained staff delivering the intervention, and an academic component will show the greatest improvement in academic achievement, though we expected that even interventions without an academic component could have some positive effect on academic achievement.

2. Method

2.1. Literature search procedures

First, we searched five different electronic reference databases for reports related to juvenile delinquency prevention programs: *ERIC*, *PsycInfo*, *Sociological Abstracts*, *Dissertation Abstracts*, and *Google Scholar*. The searches were conducted between December 2009 and June 2010 and covered studies from 1974 to June 2010. The start year of 1974 was selected due to the legislation for education reform in the provision of special education, what has become the Individuals with Disabilities Education Act (IDEA; see United States Department of Education, n.d.). The following combination of search terms was entered for all searches: "juvenile delinq*or juvenile offend*," and "intervention or rehabilitation," and "academic or achievement or grade or gpa or performance or diploma or graduation or reading or math." The first terms "delinq" and "offend" were truncated with the use of an asterisk to allow for the various words with the same beginning characters that were relevant to the search to be included (i.e. delinquency, delinquent).

There were six researchers involved in the literature search, five graduate students in a doctoral training program and one faculty member. In total 5112 document records were examined. Researchers divided the document records in a generally equal distribution of labor and examined them independently, with instructions to retrieve the record if in doubt of it being relevant. Of the documents examined, 295 were deemed potentially relevant. We obtained 250 potentially relevant documents and examined these in their entirety.

Next, we directly contacted eighteen researchers whom our database searches revealed had published two or more articles on the

topic in order to find related research that may not be included in the reference and citation databases. Similarly, significant websites relevant to the search topics, such as the OJJDP website and all related links, were examined to determine if there were additional unpublished studies relevant to the current review. We also posted a request for research on several organizations' listservs including Association for the Advancement of Behavioral and Cognitive Therapies and Trainers of School Psychologists.

Finally, we examined the references in previous syntheses of the juvenile delinquency intervention and prevention program literature, as well as the reference sections of relevant documents included in the current synthesis, to determine whether these contained mention of any reports we had not encountered through the reference database and direct-contact searches.

2.2. Criteria for including studies

For a study to be included in the research synthesis, several criteria had to be met:

(a) The target sample had to be youths who displayed some level of delinquency.
(b) The study examined the outcomes of youth who had participated in a delinquency intervention.
(c) There was some comparison group who did not participate in the intervention received no more than "treatment as usual."
(d) There was a measure of student academic achievement or school functioning.
(e) The study took place in United States or Canada after 1974.
(f) The report had to contain enough information to permit the calculation of an estimate of the relationship between participation in a juvenile delinquency intervention and the academic outcome measure.

2.3. Information retrieved from studies

Numerous characteristics of each study were included in the database. These characteristics encompassed six broad distinctions among studies. The complete coding guide is available from the first author.

2.3.1. Report characteristics

Each database entry began with the name of the author of the study, the year the study was conducted, and the type of research report (journal article, book chapter or book, dissertation, Master's thesis, private report, government report (state or federal), conference paper, or miscellaneous other type of report). We also recorded the work setting of the researchers (university, government entity, contract research firm, other), whether the research was funded and, if so, by whom (federal government, private foundation, other).

2.3.2. Research design and other study characteristics

Among other design related codes, studies were categorized into three general types of research designs. (1) Individuals could be randomly assigned to juvenile delinquency intervention programs. (2) Assignment to groups could have been accomplished using a nonrandom process but then some mechanism was used to equate individuals in the program compared to those not in the program. If equating was done, we recorded the factors used to equate groups (e.g., severity of delinquency, pretest on outcome measure, sex, ethnicity, family income level, age) and whether equating occurred as a matching procedure at the start of the study or after the study was over as a statistical process. (3) Individuals could have been assigned to groups using a nonrandom process and no attempt was made to equate the groups. For studies using a non-random process of assignment to conditions, we recorded how individuals were selected into each condition (court ordered, student characteristics, self-selection).

2.3.3. Setting and student characteristics

With regard to settings, we recorded the state and the type of community in which the study was conducted (urban, small city, suburban, rural, multiple types of communities). We also recorded the setting of the program (school, detention center, probation, home, community, hospital, residential treatment, clinic, mental health facility). With regard to student characteristic, we recorded whether an achievement label was or could be applied to the sample as a whole (gifted, average, "at risk", underachieving/below grade level, other), special diagnoses and special education status given to the sample (learning disability, receiving special education, mental illness as diagnosed by DSM-IV, English as a second language), average cognitive ability or IQ scores, and severity of delinquency. We also recorded the sample's SES, and the mean, median and range of ages, as well as the samples' gender and ethnic makeup if the group was homogeneous on these variables (only one gender or ethnicity). Although we recorded gender, ethnicity, and SES indicators, there was too much variability within samples and too few groups with homogenous samples in the reports to include these moderators in the analysis.

2.3.4. Juvenile delinquency intervention

We recorded the format of the intervention delivery (individual, groups including 2 to 10 individuals, family, classroom, whole school, or some combination of formats) and the amount of time individuals spent in the program. We recorded academic components (individualized education plan, academic skills training, GED preparation, vocational skills training, non-specific tutoring, reading-focused intervention, math-focused intervention, classroom management, and the specific academic intervention program if provided) and non-academic components (anger management training, cognitive-behavior therapy, behavioral system, family therapy, drug rehabilitation or treatment, mentorship, surveillance and monitoring, parent training, physical activity, interpersonal or relationship skills training, and the name of the specific intervention if provided) included in the intervention. We recorded the role of individuals delivering the intervention (mental health professionals, detention center staff, caseworkers or probation officers, education professionals, community volunteers, non-specified researchers). Finally, we recorded whether there was indication of treatment integrity.

2.3.5. Outcome measures

We recorded the type of each outcome measure. These included categories for: academic achievement, school enrollment, attendance, and tardiness, school attitudes, and any other academic outcome types of measures. Each type of measure was specified further using an open-ended response. In addition, achievement measures were categorized as a standardized measure, a measure developed by teachers or curriculum or textbook developers, class grades or other teacher ratings. Because we grouped all attendance related outcomes (e.g. enrollment, attendance, absenteeism, and tardiness) together for analyses, we made sure that the direction of the effects consistently indicated that a positive effect of any attendance related outcome represented a more desirable result for intervention participants. That is, although the desirability of the various measures of attendance-related phenomenon varied (i.e. attendance versus absenteeism), prior to meta-analyzing these effects, we made sure that a positive effect indicated greater attendance, less absenteeism, less tardiness, and greater enrollment among intervention participants compared to control participants. Conversely, negative effects for attendance-related outcomes relayed the opposite.

For each measure we recorded whether it was provided by the juvenile, teacher, parent, a probation or detention center officer, other school personnel or records, a probation or detention center officer and who the referent for the measure was (using the same categories). We also recorded any evidence provided about the reliability or validity of the outcome measure. Finally, we recorded when the measure was taken relative to the last day of the program in days.

As is true in all meta-analyses, many of the study characteristics, including the specific academic intervention components and psychosocial interventions, we coded were either not reported often enough or exhibited too little variability across studies to be examined as moderators of intervention effects.

2.3.6. Effect size estimation

We used the standardized mean difference to estimate the effect of the juvenile delinquency prevention program on academic outcomes. The *d*-index (Cohen, 1988) is a scale-free measure of the separation between two group means. Calculating the *d*-index for any comparison involves dividing the difference between the two group means by their average standard deviation. This calculation results in a measure of the difference between the two group means expressed in terms of their common standard deviation. Thus, a *d*-index of .25 indicates that one-quarter standard deviation separates the two means. In the meta-analysis, we subtracted the 'no program/no intervention' mean from the 'program/intervention' mean and divided the difference by their weighted average standard deviation. Thus, positive *d*-indexes indicate that the individuals in a juvenile delinquency program had better achievement or higher scores on other school-related outcome measures.

If available, we calculated effect sizes based on the means and standard deviations of the participants' outcomes. These are listed in result tables as the unadjusted effect sizes. If means and standard deviations were not available, we retrieved the information needed to calculate *d*-indexes indirectly from the inferential statistics that were reported, including p-values, t-tests, F-tests, mean gain scores, and chi-square (see Borenstein, 2009). It is worth noting that when reports that were deemed potentially relevant were eventually excluded, the most common reason was that the report did not meet some criteria in the design (i.e., inability to differentiate between control and experimental groups, or no academic outcomes reported). No studies were excluded due to the inability to calculate effect sizes if the study otherwise met the inclusion criteria. Adjusted effect sizes, to account for pre-intervention scores on the outcome measures, also were calculated if the necessary information was available in the following order of preference:

1. By calculating pre-test and post-test effect sizes separately (based on pre-test M/SD and post M/SD) and taking the difference
2. By using adjusted Ms/SDs but calculating as regular post-test ES
3. By using F-test from ANCOVA but treating it as a regular F test to calculate imprecise adjusted effect size

If the post-test scores were mean gain scores (MGSs), they were calculated as if regular post ES with Ms/SDs.

2.3.7. Coder reliability

Each research report was coded by two coders. If there was a discrepancy in coding this was first discussed by the coders. If the disagreement could not be resolved, the first or second author was consulted. Because all studies were independently coded twice and continuing disagreements were easily resolved by a third independent coder, we did not calculate reliability for this process (which would have entailed training three more coders and having them code at least a subset of studies).

2.4. Methods of data integration

First, we examined the distribution of effect sizes to determine if any were statistical outliers. Grubbs (1950) test, also called "the maximum normed residual test" was applied (see also, Barnett & Lewis, 1994). This test identifies outliers in univariate distributions and does so one observation at a time. If outliers were identified, (using *p* < .05, two-tailed, as the significance level) these values were set at the value of their next nearest neighbor.

Table 1
Characteristics of reports included in the meta-analysis.

Author (year)	Type of report	Research design	Duration of intervention	Academic program component	Program staff qualifications	Program setting	Average age of sample	Academic outcomes and effect sizes
Arbuthnot and Gordon (1986)	Journal article	Random assignment	45 min for 6–20 weeks	None specified	Doctorate	Public school	14.54	Post-intervention Achievement (grades) 0.6763 School absenteeism 0.4935 School tardiness 0.8729 Follow up School absenteeism 1.0984
Burdsal and Buel (1980)	Journal article	Non-equivalent control group with equating	3 times 2 week expeditions	Designed goals for school	Unclear	Camp setting	Unclear	Post-intervention Achievement (grades) .1317 School absences .029 Follow up: Achievement (grades) .3318 School absences .6033
Cox, Carmichael, and Dightman (1977)	Journal article	Non-equivalent control group	Unclear	Unclear	Unclear	Unclear	15.03	Post-intervention School attendance (enrolled in school) (−) .343
Davidson and Wolfred (1977)	Journal article	Non-equivalent control group with equating	Unclear	None specified	Unclear	Residential facility	14.2	Post-intervention Achievement (grades) .006 School attendance .6189 Follow up: Achievement (grades) .0122 School attendance .5151
Freiden (2005)	Dissertation	Random assignment	4 weeks	Standard school experiences	School counselor	Public school	9–12th grade	Post-intervention Achievement (grade point average) .179 School attendance (days of school missed) (−) .436
Greenwood, Deschenes, and Adams (1993)	Private report	Random assignment	Unclear	None specified	Unclear	Community	Unclear	Post-intervention School attendance (enrolled in school) (−) .5177
Newton (1994)	Dissertation	Random assignment	10–13 h	Achievement mentoring		Public school	7th–8th graders	Post-intervention Achievement (grade point average) (−) .3415
O'Dell (1980)	Journal article	Random assignment	4–12 weeks	Tutoring, vocational training, GED program	Unclear	Probation setting	15 years 10 months– 16 years 3 months	Follow up Attendance (school enrollment or employment) 3-months follow up = 2.5327 6-months follow up = 2.0256 9-months follow up = 2.3379
O'Sullivan (1979)	Dissertation	Non-equivalent control group without equating	10 sessions	School behavior check sheets	Licensed psychologists	Residential facility	Median age: 13 years range: 11–14	Post-intervention Achievement (growth score on Wide Range Achievement Test) .0643 School attendance (absences) (−) 1.935

Study	Report type	Design	Duration	Additional components		Setting	Age	Outcomes and effect sizes
Piercy and Lee (1976)	Journal article	Non-equivalent control group without equating	60 minutes per week across 10 weeks	None specified	Unclear	Probation setting	Oldest participant: 17	Post-intervention School attendance (absences) (−) 1.326
Shoji (2005)	Dissertation	Non-equivalent control group without equating	9 months	Tutoring	Unclear	Community	10–14 years of age	Time 1 post-intervention Achievement (grades) .011 (surburban sample) School attendance .028 (suburban sample) (−) .07 (urban sample) (−) .259 (rural sample) Time 2 follow up Achievement (grades) .065 (suburban) (−) .004 (urban) .005 (rural) School attendance .204 (suburban) (−) .021 (urban) (−) .075 (rural) Time 3 follow up Achievement (grades) (−).035 (suburban) (−) .198 (urban) (−) .035 (rural) School attendance .017 (suburban) (−) .021 (urban) .019 (rural)
Valentine, Griffith, Ruthazer, Gottlieb, and Keel (1998)	Journal article	Non-equivalent control group with equating	Unclear	Parent–teacher meetings and general achievement support	Unclear	Public school	12 years 9 months	Post-intervention Achievement (usual grades are C, D, or F) .4918 School attendance (cut class often) (−) .1227 Follow up Achievement (usual grades C, D, or F) .6559
Weeks (1984)	Dissertation	Random assignment	6 months	Unclear	Unclear	Probation setting	Intervention group: 15.35 years Control group: 14.83 years	Post-intervention Achievement (grades) (−) .5693 School attendance (absences) (−) .6256 Follow up Achievement (grades) (−) .264 School attendance (absences) (−) .4984
White (1979)	Dissertation	Non-equivalent control group without equating	Unclear	Vocational training	Unclear	Juvenile detention	16 years 2 months	Post-intervention Achievement (mean score on reading of Metropolitan Achievement Test) .292 Achievement (mean score on Math of Metropolitan Achievement Test) .364 School attendance (cut class often) .7244

Note: complete references for all reports are noted with an "*" in the reference section. All effect sizes in the positive direction indicate that the intervention group mean was higher on that measure compared to the control or comparison group mean score. Negative (−) effect sizes indicate the intervention group mean was lower than the comparison group mean.

2.4.1. Publication bias

Even though we used several search techniques, there is still the possibility that we did not obtain all studies that have investigated the effects of juvenile delinquency prevention programs. Therefore, we used Duval and Tweedie's (2000a, 2000b) trim-and-fill procedure to test whether the distribution of effect sizes used in the analyses was consistent with the variation in effect sizes that would be predicted if the estimates were normally distributed. If the distribution of observed effect sizes was skewed, indicating a possible bias created either by the study retrieval procedures or by data censoring on the part of authors, the trim-and-fill method provides a way to estimate the values from missing studies that need to be present to approximate a normal distribution. Then, it imputes these missing values, permitting an estimate of the impact of data censoring on the observed distribution of effect sizes.

2.4.2. Calculating average effect sizes

We used a weighted procedure to calculate average effect sizes across all comparisons to account for different size groups (Borenstein, Hedges, Higgins, & Rothstein, 2005). In this procedure, each independent effect size was first multiplied by the inverse of its variance. The sum of these products was then divided by the sum of the inverses. Also, 95% confidence intervals were calculated for average effects. If the confidence interval did not contain zero, then the null hypothesis of no difference between individuals who participated and those who did not participate in a program can be rejected.

2.4.3. Identifying independent hypothesis tests

One problem that arises in calculating effect sizes involves deciding what constitutes an independent estimate of effect. Here, we used a shifting unit of analysis approach (Cooper, 2010). In this procedure, each effect size associated with a study is first coded as if it were an independent estimate of the relationship. For example, if a single sample permitted comparisons of the effect of juvenile delinquency prevention interventions on school enrollment and daily attendance, two separate effect sizes were calculated. For estimating the overall effect of intervention on school attendance-related outcomes, these two effect sizes were averaged prior to entry into the analysis, so that the sample only contributed one effect size. However, in an analysis that examined the effect of intervention on attendance and enrollment separately, this sample would contribute one effect size to each estimate of a category mean effect size.

The approach retains as much data as possible from each study while holding to a minimum any violations of the assumption that data points are independent. Also, because effect sizes are weighted by sample size in the calculation of averages a study with many independent samples containing just a few participants will not have a larger impact on average effect size values than a study with only a single, or a few, large independent samples.

2.4.4. Tests for moderators of effects

Possible moderators of the effects of juvenile delinquency prevention program were tested using homogeneity analyses (Cooper, Hedges, & Valentine, 2009). Homogeneity analyses compare the amount of variance in an observed set of effect sizes with the amount of variance that would be expected by sampling error alone. The analyses can be carried out to determine whether (a) the variance in a group of individual effect sizes varies more than predicted by sampling error or (b) a group of average effect sizes varies more than predicted by sampling error. In the latter case, the strategy is analogous to testing for group mean differences in an analysis of variance or linear effects in a multiple regression.

2.4.5. Fixed and random effects

When an effect size is said to be "fixed" it assumes that error is due solely to differences among participants sampled in the study. However, it is also possible to view studies as containing other random influences, including differences in treatment providers, facilities, community economics, and so on. This view assumes that programs, centers, schools, or even whole communities in our meta-analysis also constitute a random sample drawn from a (vaguely defined) population of conditions under which juvenile delinquency intervention programs take place. If it is believed that such random variation in programs is a significant component of error, a random effect model should be used that takes into account this study-level variance in effect sizes (see Hedges & Vevea, 1998, for a discussion of fixed and random effects). Given the diversity of settings, the inherent variability in youths who engage in delinquency as a broad construct, and the wide variety of services offered within the context of most juvenile delinquency programs, we used a random effects models.

2.4.6. Software

All statistical analyses were conducted using the Comprehensive Meta-Analysis statistical software package (Borenstein et al., 2005). Only outcome measures that were collected in studies described in at least four separate reports with at least four independent samples had meta-analytic procedures applied to them.

3. Results

The literature search located a total of 15 reports that compared youths who attended delinquency prevention program to youths who did not attend any delinquency program on some measure of school functioning. The 15 reports provided 134 separate effect sizes (92 unadjusted and 42 adjusted) based on 20 separate samples. The 15 reports appeared between 1976 and 2005. The sample sizes ranged from 19 to 417. Eight of these reports were published. The characteristics of the studies are available in Table 1.

Thirteen of the 15 reports contained 18 separate samples with information that allowed juveniles who participated in a program to be compared to juveniles who did not without statistical adjustments to the comparison. Of these, 10 reports and 12 samples also provided a comparison of juveniles who participated in a program to those who did not that was statistically adjusted post hoc by a pre-test measure of the outcome to make the groups more comparable. There were an additional 2 reports containing 2 samples for which only an adjusted comparison could be obtained. Thus, a total of 12 reports containing 14 samples provided an effect adjusted by a pre-test measure of the outcome to make the groups more comparable.

Of these 15 reports, 10 reports and 14 separate samples provided information on academic achievement. The 10 reports provided 33 separate effect sizes. Five of these samples provided only unadjusted effect of program on academic achievement, two provided only effects statistically adjusted by a pre-test measure of the achievement to make the groups more equivalent, and seven samples provided both an unadjusted and adjusted effects of juvenile delinquency prevention program.

Seven reports that assessed achievement and four additional reports (including 15 independent samples total) provided information on school attendance, enrollment, and tardiness. The eight reports provided 40 separate effect sizes. Eight of these samples provided only unadjusted effects, two samples provided only effects adjusted by a pre-test measure of the outcome, and five samples provided both an unadjusted comparison as well as a comparison statistically adjusted for a pre-test measure of the outcome to make the groups more equivalent.

One report that assessed all previous outcomes and one additional report (including 3 independent samples total) provided information on maladaptive school attitudes. The three samples provided 10 separate effect sizes. All three of these samples provided both an unadjusted comparison as well as a comparison statistically adjusted for a pre-test measure of the outcome to make the groups more equivalent.

3.1. Overall effects of juvenile delinquency intervention program on school functioning

Table 2 presents the results of the analyses examining the overall effect of juvenile delinquency intervention program using achievement and related school outcomes as the outcome variable. For each outcome, Grubb's test was used to identify outliers within that set of effect sizes. No outliers were identified among the achievement, attendance, or school attitudes effect sizes. Results are reported for random-error models only.

3.1.1. Achievement

The unadjusted effects of program on academic achievement varied between $d = -.57$ and $+.66$. Of the 21 unadjusted (non-independent) effect sizes assessing the effect of program on achievement, 11 were in a positive direction, and 10 in a negative direction. The weighted average d-index across all samples was .12 and was not significantly different from 0 (95% CI = $-.06/.30$). The tests of the distribution of d-indexes revealed that we could reject the hypothesis that the effects varied only due to sampling error and therefore were estimating the same underlying population value, $Q(11) = 53.27$, $p<.001$.[1]

The adjusted effects of program on achievement varied between $d = -.48$ and $+1.12$. Of the 12 adjusted (non-independent) effect sizes assessing the effect of program on achievement, 6 were in a positive direction, and 6 in a negative direction. The weighted average d-index was .02 and was not significantly different from 0 (95% CI = $-.18/.22$). Again, the tests of the distribution of d-indexes revealed that we could reject the hypothesis that the effects were estimating the same underlying population value, $Q(8) = 17.19$, $p = .03$.

Trim-and-fill analyses were conducted in several different ways. First, we performed the analyses looking for asymmetry (see, Borenstein et al., 2005) while searching for possible missing effects on the left side of the distribution (those that would reduce the size of the positive average d). For the unadjusted achievement effects, three additional effect sizes were imputed. This would change the mean effect $d = -.02$ (95% CI = $-.23/.19$). No additional effect sizes were imputed on the left side of the distribution. For the adjusted achievement effects, no additional effect sizes were imputed on the left side of the distribution.

Trim-and-fill analyses conducted to search for possible missing effects on the right side of the distribution (those that would increase the size of the positive average d) revealed that one additional effect size was imputed among the adjusted effects. Imputing this value would change the mean effect to $d = .07$ (95% CI = $-.14/.27$). No additional effect sizes were imputed on the right side of the distribution. For the unadjusted achievement effects, no additional effect sizes on the right side of the distribution were imputed. Thus, when testing for possible data censoring, the effect of juvenile delinquency intervention program on achievement may not be significantly different from zero.

3.1.2. School Attendance Related Outcomes

The unadjusted effects of program on attendance, enrollment, and tardiness varied between $d = -.78$ and $+1.94$. Of the 28 unadjusted

Table 2
Overall effects of juvenile delinquency program on school related outcomes.

Outcome	Unadjusted d-indexes				Adjusted d-indexes			
	k	d	$-/+$ CI	Q	k	d	$-/+$ CI	Q
Achievement	12	.12	$-.06/.30$	53.27***	9	.02	$-.18/.22$	17.19*
Attendance	13	.11	$-.13/.35$	140.83***	7	$-.14$	$-.58/.29$	58.18***
Maladaptive school attitudes	3	.19	$-.08/.46$	8.72**	3	$-.12$	$-.34/.11$	6.03+

*** $p<.001$.
  * $p<.05$.
 ** $p<.01$.
  + $p<.10$.

(non-independent) effect sizes assessing the effect of program on attendance related outcomes, 16 were in a positive direction, and 12 in a negative direction. The weighted average d-index was .11 and was not significantly different from 0 (95% CI = $-.13/.35$). The tests of the distribution of d-indexes revealed that we could reject the hypothesis that the effects varied only due to sampling error and therefore were estimating the same underlying population value, $Q(12) = 140.83$, $p<.001$.

The adjusted effects of program on attendance related outcomes varied between $d = -1.10$ and $+.78$. Of the 12 adjusted (non-independent) effect sizes assessing the effect of program on attendance related outcomes, 5 were in a positive direction, and 7 in a negative direction. The weighted average d-index was $-.14$ and was not significantly different from 0 (95% CI = $-.58/.29$). Again, the tests of the distribution of d-indexes revealed that we could reject the hypothesis that the effects were estimating the same underlying population value, $Q(6) = 58.18$, $p<.001$.

For the unadjusted attendance related effects, trim and fill analyses revealed that three additional effect sizes were imputed on the left side of the distribution. Imputing these values would change the mean effect to $d = -.10$ (95% CI = $-.37/.16$). No additional effect sizes were imputed on the left side of the distribution among the unadjusted or adjusted effects. Likewise, no additional effect sizes were imputed on the right side of the distribution among either unadjusted or adjusted effects.

3.1.3. Maladaptive school attitudes

The unadjusted effects of program on maladaptive school attitudes varied between $d = -.17$ and $+.48$. Of the 5 unadjusted (non-independent) effect sizes assessing the effect of program on maladaptive school attitudes, 4 were in a positive direction, and 1 was in a negative direction. The weighted average d-index was .19 and was not significantly different from 0 (95% CI = $-.08/.46$). The tests of the distribution of d-indexes revealed that we could reject the hypothesis that the effects varied only due to sampling error and therefore were estimating the same underlying population value, $Q(2) = 8.72$, $p<.01$.

The adjusted effects of program on school attitudes varied between $d = -.54$ and $+.13$. Of the five adjusted (non-independent) effect sizes assessing the effect of program on school attitudes, two of them were in a positive direction, and three in a negative direction. The weighted average d-index was $-.12$ and was not significantly different from 0 (95% CI = $-.34/.11$). The tests of the distribution of d-indexes revealed that we could not reject the hypothesis that the effects were estimating the same underlying population value, $Q(2) = 6.03$, $p = .05$.

Trim-and-fill analyses for the unadjusted school attitude effects revealed that no additional effect sizes were imputed on the left side of the distribution. One additional unadjusted effect was imputed on the right side of the distribution. Imputing these values would change the mean effect to $d = .15$ (95% CI = $-.03/.32$). For the adjusted school attitude effects, no additional effect sizes were imputed on the left or right side of the distribution.

---

[1] As a note of clarification, in line with the shifting-unit-of-analysis approach, every relevant effect size that a sample provided was coded. However, prior to calculating a weighted average effect, non-independent effects were averaged prior to entry into the analysis, so that each sample only contributed one effect size. Using the analysis of the overall effect of achievement as an example, although 21 separate effects sizes relevant to the intervention-achievement relationship could be retrieved, some of these effects came from the same sample and were subsequently averaged prior to computing the weighted average effect of intervention on achievement. Consequently, the number of independent samples contributing to the overall effect was 12 and the degrees of freedom for testing homogeneity was 11. This situation applied to all subsequent tests of the overall effects of delinquency intervention on the major categories of school outcomes.

3.2. Moderators of the effects of juvenile delinquency program

Next, we conducted moderator analyses of the juvenile delinquency program on academic achievement and attendance related outcomes using eight moderators of methodological and theoretical interest for academic achievement and nine moderators for attendance. Given the limited number of samples contributing to the database of effects for each outcome, we attempted to limit the number of moderators examined to those we felt would be of greatest practical and theoretical interest. Moderator analyses were conducted using the unadjusted effects only since fewer than ten samples provided adjusted effects for each of the relevant outcomes. Likewise, no moderator analyses were conducted for program effects on maladaptive school attitudes due to the limited number of contributing samples. Tables 3 and 4 present the details of these results and provide sample numbers, effect sizes, confidence intervals, and Q values for each moderator test. Below, we briefly summarize these analyses narratively and provide information not in the tables when appropriate.

3.2.1. Methodological moderators of achievement

We conducted analyses exploring two moderators of the juvenile delinquency program–achievement association related to the methodology of the studies: (a) the type of report and (b) the research design, that is, whether or not the study used a procedure in which participants were randomly assigned to conditions or not and if not, whether or not a matching procedure was used to equate individuals across groups. The unadjusted achievement effect size was significantly greater among published studies than unpublished studies.

Table 3
Moderators of the unadjusted effect of juvenile delinquency program on academic achievement.

Moderator	k	d	−/+ CI	Q
Type of report	12			9.11[*]
Published	5	.39	.15/.63	
Unpublished	7	−.03	−.16/.10	
Research design	12			14.93[**]
Nonequivalent control group without matching	3	.44	.14/.74	
Nonequivalent control group with matching	3	.28	.03/.53	
Randomized experiment	6	−.07	−.16/.02	
Intervention time	8			.82
<6 months	4	.07	−.21/.34	
6 to 12 months	4	−.07	−.16/.03	
Academic component	9			.04
No	4	.02	−.31/.34	
Yes	5	−.02	−.15/.11	
Intervention staff	5			.17
Mental health professionals	3	.19	−.13/.51	
Other	2	.04	−.61/.67	
Setting	9			.02
Non-school	7	−.01	−.14/.11	
School	2	−.06	−.56/.45	
Average sample age	9			.20
<15	7	−.05	−.14/.04	
15 to 18	2	.05	−.39/.49	
Delinquency severity	9			.00
More severe (legal action taken)	2	−.02	−.75/.70	
Less severe (no legal action taken)	7	−.04	−.13/.04	

Note that for all moderators except publication status and research design, only randomized experiments and nonequivalent control group designs with matching were included (9 samples). Also note that the number of samples contributing from one moderator analysis to the next varies and that the number of samples for individual moderator analyses is often less than the number of samples contributing to overall effects of program. This was because there were some reports that did not provide information for various characteristics of interest or characteristics of the sample did not allow for it to be conceptually classified among categories of interest for the moderator analyses. When this was the case, the effects for such samples were excluded from the individual moderator analysis.

 * $p < .01$.
 ** $p < .001$.

Table 4
Moderators of the unadjusted effect of juvenile delinquency program on school attendance outcomes.

Moderator	k	d	−/+ CI	Q
Type of report	13			.00
Published	5	.09	−.54/.71	
Unpublished	8	.08	−.14/.29	
Research design	13			.49
Nonequivalent control group without matching	3	.27	−.66/1.19	
Nonequivalent control group with matching	2	−.13	−.98/.73	
Randomized experiment	8	.17	−.09/.44	
Intervention Time	10			.28
<6 months	6	.16	−.34/.66	
6 to 12 months	4	.02	−.16/.19	
Academic component	12			.32
No	4	.15	−.43/.72	
Yes	8	.34	.01/.67	
Intervention staff	5			.42
Mental health professionals	3	.33	.12/.53	
Other	2	.07	−.69/.83	
Setting	10			1.60
Non-school	9	.09	−.16/.34	
School	1	.44	−.04/.91	
Average sample age	8			7.25[*]
<15	5	−.06	−.27/.14	
15 to 18	3	.80	.21/1.38	
Delinquency severity	10			.08
More severe (legal action taken)	5	.16	−.49/.82	
Less severe (no legal action taken)	5	.07	−.10/.23	
Outcome type	10			.18
Attendance	7	.04	−.16/.24	
Enrollment	3	.30	−.88/1.48	

Note that for all moderators except publication status and research design, only randomized experiments and nonequivalent control group designs with matching were included (9 samples). Also note that the number of samples contributing to from one moderator analysis to the next varies and that the number of samples for individual moderator analyses is often less than the number of samples contributing to overall effects of program. This was because there were some reports that did not provide information for various characteristics of interest or characteristics of the sample did not allow for it to be conceptually classified among categories of interest for the moderator analyses. When this was the case, the effects for such samples were excluded from the individual moderator analysis.

 * $p < .01$.

The association between the magnitude of achievement effect sizes and whether or not the study used randomized assignment or a matching procedure prior to conducting the study was examined. Samples were considered to have employed a matching procedure if they matched individual participants of the juvenile delinquency prevention program to individuals not in the program on various characteristics. Studies in which matching was conducted at the site, school, district, or community level were not considered to have employed matching for the purposes of this analysis.

The average unadjusted effect on achievement varied by research design. We then proceeded to conduct contrasts to compare the unadjusted effect of program on achievement for each pair of research designs. The average unadjusted effect of program from randomized experiments was significantly smaller than the unadjusted effect of nonequivalent control group designs with matching, $(Q(1) = 6.54, p < .02)$, as well as than the unadjusted effect of non-equivalent control group designs without matching, $(Q(1) = 9.82, p < .01)$. There was no statistically significant difference between non-equivalent control group designs with versus without matching $(Q(1) = 2.02, p = .15)$, though the effect among non-equivalent control group designs without matching was larger.

Based on both a theoretical expectation that a non-equivalent control group design without matching would produce an unacceptably biased assessment of program unadjusted effects, as well as the finding of this moderator analysis that suggested that this weakest design indeed demonstrated an inflated unadjusted effect compared to randomized experiments, non-equivalent control group designs without matching were excluded from further analyses examining theoretical moderators

of unadjusted effects. We decided to retain non-equivalent control group designs with matching for further moderator analyses of unadjusted effects based on the assumption that this design provides a reasonable assessment of program effects due to the attempts to equate groups on various characteristics.

3.2.2. Treatment and sample moderators of achievement

We conducted analyses exploring six moderators of the juvenile delinquency program–achievement association related to the characteristics of the treatment and of the juveniles: (a) duration of the intervention, (b) presence of an academic program component, (c) program staff qualifications, (d) program setting, (e) average age of the sample, and (f) severity of sample delinquency. While we would have liked to examine whether the program–achievement association varied by the type of outcome (e.g. grades versus standardized test scores), we were unable to because there was little variability among samples on this characteristic. Achievement was assessed primarily as course grades, and it was rare that any study used standardized measures of achievement. None of the treatment or sample characteristics was found to moderate the effect of program on achievement.

3.2.3. Methodological moderators of school attendance related outcomes

We next conducted analyses exploring methodological moderators of the juvenile delinquency program–attendance association. The average weighted effect of program on school attendance related outcomes did not significantly vary by type of report (published versus unpublished) or research design.

Although our moderator analysis indicated that there was not a significant variability in the effects across research designs, to be consistent with moderator analyses of the program–achievement association, non-equivalent control group designs without matching were excluded from further analyses examining theoretical moderators of unadjusted effects on attendance related outcomes.

3.2.4. Treatment and sample moderators of school attendance related outcomes

We conducted analyses exploring seven moderators of the delinquency intervention and school attendance association related to several characteristics: (a) duration of the intervention, (b) presence of an academic program component, (c) program staff qualifications, (d) program setting, (e) average age of the sample, (f) severity of sample delinquency, and (g) type of outcome.

The unadjusted effect of intervention on school attendance related outcomes was found to significantly vary depending on the average age of the sample. The intervention had a significant positive unadjusted effect on school attendance related outcomes among samples with an average age between 15 and 18, but no significant effect among samples with an average age below 15 years. The effect of intervention on school attendance related outcomes did not vary by the duration of the intervention, presence of an academic component, program setting, severity of delinquency or who administered the intervention.

4. Discussion

The goal of this meta-analysis was to synthesize the current state of the literature on the effects of juvenile delinquency intervention programs on academic achievement and related indicators of school success, including school attendance and attitudes toward school. Results highlight several important points about the literature on delinquency interventions and academic outcomes. The most important point is how clearly this investigation highlights the gaps and lack of trustworthy information in the current literature.

The overall results, although unclear in many ways due to research design concerns, suggest that juvenile delinquency interventions are generally ineffective in improving academic outcomes. This was the

case even when there was an academic component in the delinquency program. Further, the tests of methodological moderators are consistent with results from prior meta-analyses (see Lipsey & Cullen, 2007): published studies and weaker research designs tended to reveal larger effects. More specifically, while small to moderately sized positive effects of intervention on achievement were found among studies using non-equivalent control groups, these designs do not typically allow for as trustworthy of conclusions as those coming from randomized experiments. For the most trustworthy research designs, the effect of intervention on achievement was not significant.

Given the known links between delinquency and academic achievement, we were somewhat surprised at the limited number of interventions that examined academic outcomes. While we initially found over 250 reports that seemed relevant to the topic of interest, after sorting through studies, we found only 15 reports that met our inclusion criteria. Thus, it would seem that the effect of juvenile delinquency interventions on academic outcomes is indeed under-investigated area of research. In addition, the most recent study that met our criteria was reported in 2005, so the most recent 5 years yielded no reports, which is also indicative of a concerning lack of investigations about this important topic.

Another limitation of the available literature relates to how scholars reported information about the studies. Though we had thought it was plausible that a number of characteristics of interventions and samples might explain variability in effects, the small population of studies that were eventually included in the synthesis and the limited information provided about intervention and sample characteristics in reports prohibited us from conducting more nuanced (and potentially, more informative) moderator analyses. For example, we would have liked to have compared the effect of intervention on academic outcomes for juvenile delinquents of varying race, sex, achievement levels, or educational needs, but there was limited reporting and limited variability among samples on these characteristics. Likewise, we would have liked to compare interventions coming from various theoretical traditions or containing various types of academic components (rather than just dichotomize interventions that did versus did not have any type of academic component), but often studies did not provide much information on the academic components included.

In addition, as is typical in most juvenile delinquency intervention research (Lipsey & Cullen, 2007), theory-driven or manualized programs seemed to be the exception. We examined reports for any information about a theory-driven or empirically evaluated educational component, and there were no studies that included manualized or rigorously tested educational interventions. In other words, juvenile delinquency interventions are weak in terms of using best practice in educational and psychosocial intervention approaches, which may explain the overall lack of effectiveness even in the well-designed studies.

With those limitations in mind, the findings from this meta-analysis did show that after excluding the least trustworthy studies (non-equivalent control groups without equating), juvenile delinquency interventions may have a positive effect on school attendance among older delinquents (ages 15 to 18) and the effect of intervention among this group was significantly different from that on younger delinquents. There are two important considerations to understand the implications of this finding. Although this finding is consistent with the Lipsey (1999) synthesis showing age was positively associated with a better response to the delinquency intervention, this finding here is based on unadjusted effects, so conclusions are tentative. In the disruptive behavior disorders literature, which overlaps with the general juvenile delinquency literature, children who display more symptoms and higher levels of problem behaviors at younger ages are often simply more distressed and have a poorer overall prognosis (Granic & Patterson, 2006; Murray-Close et al., 2010; Wiesner, Kim, & Capaldi, 2005). Children who display symptoms in milder form or who display delinquency at older ages may represent the "milder" form of the disruptive behavior

disorders. This age phenomenon may explain the higher response to intervention overall. The finding here is about school attendance, not academic skill, so it is unclear if the attendance would ultimately improve academic achievement based on current available studies. So, this particular finding is exploratory because it is not only the age and attending school that is of interest, but also the educational skill level that is important to address. The available reports did not document educational skills well, so the competing reasons for this tentative finding stand to be clarified in future research.

Next, much like prior meta-analyses on the effect of juvenile delinquency interventions on recidivism outcomes (Lipsey & Cullen, 2007), heterogeneity of the samples and a lack of reporting the information separately for groups with different characteristics (e.g. race, sex, achievement levels, etc…) made it challenging to draw meaningful conclusions about the effects of delinquency interventions on academic outcomes. There was a high possibility that there might be variability in outcomes as a function of such sample characteristics rather than the intervention. Overall, this investigation highlights the continued need for higher quality research designs, analyses separated for specific groups included in an intervention, and reporting more information about the population served.

Central to the juvenile delinquency literature is the issue of crime, but how schools and education relate to the question of 'what works, for whom, and under what conditions' remains an important issue. As offered by Lipsey and Cullen (2007), the best treatment is that which is based on determining the needs of the youth, particularly the dynamic needs as opposed to static background characteristics. Lipsey and Cullen (2007) articulated the need for theory-driven approaches, and particularly noted the potential utility of the perspective offered by Andrews and colleagues (see Andrews, 1995) regarding the principle of responsivity. This principle posits the idea that the malleable risks associated with crime can be identified and addressed, or matched with specific intervention components in response to the targeted identified need of the offender (Lipsey & Cullen). Given the complex needs in juvenile delinquency populations, the identified needs would typically include emotional, behavioral and academic areas. Then, based on those needs, provide treatment. When needs are matched to the treatment offered, the effects are likely to be larger and more positive in terms of juvenile offending (Andrews & Bonta, 2006).

In addition, achievement does not progress in a linear fashion without appropriate learning experiences, and the gap between expected achievement and actual grade level is quite important in understanding school disengagement (Simpson, Swanson, & Kunkel, 1992). There was no study that included a program to match needs for academic outcomes based on empirical or rigorously tested academic programs. For example, in other intervention literature, including reading interventions and psychosocial interventions, the targeted needs of the intervention recipients are matched with the intervention components to produce the highest positive effects (Edmonds et al., 2009; McCarty & Weisz, 2007). This notable gap is an important area to address in determining what could work for improving educational outcomes among juvenile delinquents, and how meeting educational needs could relate to lower recidivism.

4.1. Future directions

In general, the most important direction we have for future research is to rely on study designs that provide trustworthy examinations of intervention effectiveness. More specifically, using randomized controlled experimental designs is strongly encouraged. A next-best option is to increase the use of equivalent group designs with matching and statistical post-hoc controls. The use of non-equivalent groups without equating procedures is not recommended as it does not allow for trustworthy conclusions about program effects or differences in effects across groups.

In addition, information about academic skills and intervention characteristics in examinations of juvenile delinquency intervention effectiveness is sorely needed. For example, in reading interventions there are differences in academic effects based on the specific type of intervention offered, such as tutoring compared to a reading fluency intervention (Edmonds et al., 2009). This is important in psychosocial and educational research. As highlighted by Weisz et al. (2006) in a meta-analysis of psychosocial interventions with youth, specificity of the intervention itself is very important to address. This includes the type of intervention, dosage or amount of exposure to the intervention, and a clear description of the actual intervention components. The current synthesis of research could not include these moderators due to the lack of specificity reported about the level of academic performance, the type of interventions, or lack of specific intervention altogether. Future researchers are encouraged to (a) report the academic skill levels of their population, (b) include specific interventions and also (b) describe the intervention in detail to allow for more clarity about what works.

Similar to other scholars who focused exclusively on school settings (Wilson & Lipsey, 2007) or corrections/detention settings (Krezmien & Mulcahy, 2008; Leone et al., 2005), our study highlights the need for educational interventions to be incorporated into delinquency programs regardless of the setting. Specifically, implementing and examining effects of the interventions listed on the public database maintained by the United States Department of Education and Institute of Education Sciences, the *What Works Clearinghouse (WWC; n.d.)*, a resource for educational interventions based on rigorous empirical evaluation standards, is recommended as a next step in juvenile delinquency research. Given these recommendations, partnerships between researchers in fields of psychology and education and juvenile facilities remain necessary and are highly encouraged.

We would also highlight that academic achievement outcome variables that were most often used (grades) may not be the best indicator of skill improvements resulting from an intervention. Thus, future research might consider using more sensitive measures that will increase the likelihood of detecting changes in academic skills due to the intervention if they are present. Using standardized measures of reading skill is also highly recommended (Edmonds et al., 2009). Examples of these include the *Woodcock-Johnson IIIr Diagnostic Reading Battery* (WJIIIr-DRB; Schrank, Mather, & Woodcock, 2004), the *Gray Oral Reading Test, 4th edition* (GORT4; Wiederholt, J. & Bryant, B. 2001), or the *Wide Range Achievement Test, 4th edition, Progress Monitoring Version* (WRAT4-PMV; Roid, Ledbetter, & Messer, 2006). In addition, the *Response to Intervention* academic/instructional literature contains numerous tools for measuring academic skill improvement using reliable and valid measures that are also sensitive to changes in skill level. There are several readily available tools to implement at the classroom or individual student level that could be easily adapted to a variety of juvenile delinquency program settings (see Busch & Reschly, 2007; Daly, Martens, Barnett, Witt, & Olson, 2007).

5. Limitations

Typical of many meta-analyses, there are limitations. We expended considerable time and energy locating reports from published and unpublished sources, but it is likely we were unable to locate every possible relevant report. Also, with the very nature of meta-analyses, the study findings are limited by the information within the original research report. Although inflated effect sizes from weaker research designs (rather than from program effects) are often reported in meta-analytic studies, the studies included here seem to be on par with the typical effect sizes in well-controlled studies in juvenile justice literature in general (Lipsey & Cullen, 2007). Even so, our conclusions are tentative based on the quality of the data reported, the heterogeneity of the overall population, and the unique challenges of how interventions with youths engaged in delinquency

are conducted. Controlled experiments with this population are rare (but nonetheless encouraged), and quasi-experimental designs would be expected — but even quasi-experimental designs with specific academic information are sparse, and this limits the understanding of effects of the intervention on academic outcomes.

6. Conclusions

This synthesis provides a marker of the state of the literature, and there is considerable room for increased knowledge to inform juvenile delinquency research in some meaningful ways. Given that educational variables are malleable and dynamic (Edmonds et al., 2009), and are closely associated with criminal behavior (Cottle et al., 2001; Zabel & Nigro, 2007), incorporating educational aspects into rehabilitation programs seems intuitive in terms of a needs-response or responsivity principle (Andrews, 1995). It is clear that rehabilitation programs, particularly those based on specific needs of the offender, are helpful in reducing criminal activity (Lipsey & Cullen, 2007) Moreover, it seems likely that rehabilitation programs might be most advantageous when educational outcomes are also improved. Unfortunately, the effectiveness of those programs that have been tested for improving academic outcomes seems to be lackluster, or at least the evidence is lacking. A possible avenue for future research is to examine the effects of juvenile delinquency programs that make use of the recently expanded variety of programs and resources from the educational intervention literature that have been found to be most effective. The relationship between low academic achievement and delinquency is well established (Cottle et al., 2001), but the field needs to move on to exploring 'how' and 'why' this link is present in order to improve and inform future delinquency interventions and interrupt the school to prison pipeline. It would be helpful for juvenile justice and educational researchers to draw on interdisciplinary approaches and collaborate in order to address the ongoing and related problems of juvenile delinquency and school failure.

References

Andrews, D. A. (1995). The psychology of criminal conduct and effective treatment. In J. McGuire (Ed.), *What works: Reducing reoffending—Guidelines from research and practice* (pp. 35–62). New York: John Wiley.

Andrews, D. A., & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Newark, NJ: Lexis/Nexis.

*Arbuthnot, J., & Gordon, D. (1986, April). Behavioral and cognitive effects of a moral reasoning development intervention for high-risk behavior-disordered adolescents. *Journal of Consulting and Clinical Psychology*, *54*(2), 208–216.

Barnett, V., & Lewis, T. (1994). *Outliers in statistical data* (2nd ed.). New York: Wiley Press.

Borenstein, M. (2009). Effect sizes for studies with continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 221–352). (2nd ed.). New York: Russell Sage Foundation.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis (version 2) [computer software]*. Englewood, NJ: Biostat.

*Burdsal, C., & Buel, C. (1980). A short term community based early stage intervention program for behavior problem youth. *Journal of Clinical Psychology*, *36*, 226–241.

Busch, T. W., & Reschly, A. L. (2007). Progress monitoring in reading: Using curriculum-based measurement in a response-to-intervention model. *Assessment for Effective Intervention*, *32*(4), 223–230. http://dx.doi.org/10.1177/15345084070320040401.

Campbell, D. M., Marsh, D. T., & Stickel, K. (1993). Intervention with adjudicated adolescents: Academic and psychological effects. *Journal of Offender Rehabilitation*, *19* 101–111.1993-35068-001.

Christle, C. A., Jolivette, K., & Nelson, C. M. (2005). Breaking the school to prison pipeline: Identifying school risk and protective factors for youth delinquency. *Exceptionality*, *13*(2), 69–88.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cooper, H. M. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Los Angeles: Sage Publications.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York: Russell Sage Foundation.

Cottle, C. C., Lee, R. L., & Heilbrun, K. (2001). The prediction of criminal recidivism in juveniles. *Criminal Justice and Behavior*, *28*, 367–394.

Council of State Governments and Public Policy Initiative (2011). Breaking school rules: A statewide study of how school discipline relates to students' success and

juvenile justice involvement. Accessed July 26, 2011 at: http://knowledgecenter. csg.org/drupal/system/files/Breaking_School_Rules.pdf

*Cox, G. B., Carmichael, S. J., & Dightman, C. (1977). An evaluation of a community based diagnostic program for juvenile offenders. *Juvenile Justice Journal*, *28*(3), 33–41.

Daly, E., Martens, B. K., Barnett, D., Witt, J. C., & Olson, S. C. (2007). Varying intervention delivery in response to intervention: Confronting and resolving challenges with measurement, instruction, and intensity. *School Psychology Review*, *36*(4), 562–581.

*Davidson, W. S., & Wolfred, T. R. (1977). Evaluation of a community-based behavior modification program for prevention of delinquency: The failure of success. *Community Mental Health Journal*, *13*(4), 296–306.

Drakeford, W. (2002). The impact of an intensive program to increase the literacy skills of incarcerated youth. *Journal of Correctional Education*, *53*, 139–144.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405–432. http://dx.doi.org/10.1111/j.1467-8624.2010.01564.x.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455–463.

Duval, S., & Tweedie, R. (2000). A nonparametric "Trim and Fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, *95*, 89–98.

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K., & Schnakenberg, J. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research*, *79*(1), 262–300. http://dx.doi.org/10.3102/0034654308325998.

*Freiden, J. (2005). Game: A clinical intervention to reduce adolescent violence in schools. Doctoral Dissertation. Retrieved from ProQuest Dissertation and Theses Database. (UMI Number: 3199466).

Granic, I., & Patterson, G. (2006). Toward a comprehensive model of antisocial development: A dynamic systems approach. *Psychological Review*, *113*(1), 101–131. http://dx.doi.org/10.1037/0033-295X.113.1.101.

*Greenwood, P. W., Deschenes, E. P., & Adams, J. (1993). *Chronic juvenile offenders: Final results from the skillman aftercare experience.* : RAND Corporation.

Gresham, F. M. (2007). Evolution of the response-to-intervention concept: Empirical foundations and recent developments. *Handbook of Response to Intervention*, *1*, 10–24.

Grubbs, F. E. (1950). Sample criteria for testing outlying observations. *Annals of Mathematical Statistics*, *21*(1), 27–58.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, *3*(4), 486–504.

Howden, L. M., & Meyer, J. A. (2011). *Age and sex composition: 2010: 2010 Census Briefs.* : United States Census Bureau Accessed August 10, 2011 at http://www.census. gov/prod/cen2010/briefs/c2010br-03.pdf

Keith, J., & McCray, A. (2002). What about the girls? Delinquency programming as if gender mattered. *Corrections Today*, *63*(1), 38–44.

Krezmien, M. P., & Mulcahy, C. A. (2008). Detained and committed youth: Examining differences in achievement, mental health needs, and special education status. *Education and Treatment of Children*, *31*(4), 445–464.

Kroesbergen, E. H., & Van Luit, J. H. (2003). Mathematics interventions for children with special educational needs: A meta-analysis. *Remedial and Special Education*, *24*(2), 97–114. http://dx.doi.org/10.1177/07419325030240020501.

Leone, P. E., Christle, C. A., Nelson, C. M., Skiba, R., Frey, A., & Jolivette, K. (2003). *School failure, race, and disability: Promoting positive outcomes, decreasing vulnerabilities for involvement with the juvenile delinquency system.* Retrieved May 1, 2011 from. : EDJJ: The National Center on Education, Disability, and Juvenile Justice http://www.edjj.org/Publications/ list/leone_et_al-2003.pdf

Leone, P. E., Krezmien, M., Mason, L., & Meisel, S. M. (2005). Organizing and delivering empirically based literacy instruction to incarcerated youth. *Exceptionality*, *13*(2), 89–102.

Lipsey, M. W. (1999). Can rehabilitative programs reduce the recidivism of juvenile offenders? An inquiry into the effectiveness of practical programs. *Virginia Journal of Social Policy & the Law*, *6*(3), 611–641.

Lipsey, M. W. (2009). The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims and Offenders*, *4*, 124–147.

Lipsey, M. W., & Cullen, F. T. (2007). The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annual Review of Law and Social Science*, *3*, 297–320.

Malmgren, K. W., & Leone, P. E. (2000). Effects of a short-term auxiliary reading program on the reading skills of incarcerated youth. *Education and Treatment of Children*, *23*, 239–247 2000-16730-003.

McCarty, C. A., & Weisz, J. R. (2007). Effects of psychotherapy for depression in children and adolescents: What we can (and can't) learn from meta-analysis and component profiling. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*(7), 879–886. http://dx.doi.org/10.1097/chi.0b013e31805467b3.

Murray-Close, D., Hoza, B., Hinshaw, S., Arnold, L., Swanson, J., Jensen, P., et al. (2010). Developmental processes in peer problems of children with attention-deficit/hyperactivity disorder in The Multimodal Treatment Study of Children with ADHD: Developmental cascades and vicious cycles. *Development and Psychopathology*, *22*(4), 785–802. http://dx.doi.org/10.1017/S0954579410000465.

National Center for Education Statistics (2010). *The condition of education 2010.* : U.S. Department of Education, Institute of Education Science Accessed April 2, 2011 at http://nces.ed.gov/pubs2010/2010028_2.pdf

*Newton, F. R. (1994). A study of the effectiveness of using collegiate mentors to reduce violent behavior, improve self-concept, and increase academic success in an urban middle school. Doctoral Dissertation. Retrieved from ProQuest Dissertation and Theses Database. (Order Number: 9429678).

*O'Dell, B. N. (1980). Accelerating entry into the opportunity structure: A sociologically-based treatment for delinquent youth. *Sociology and Social Research*, 312–317.

*O'Sullivan, T. (1979). The effect of short-term behavioral group counseling on academic achievement and behavior with adolescent boys in a residential treatment program. Unpublished Doctoral Dissertation. St. John's University, New York.

Office of Juvenile Justice and Delinquency Prevention (2008). *Census of juveniles in residential placement 2006.* Washington, D.C.: OJJDP Accessed April 2, 2011 at http://ojjdp.ncjrs.gov/ojstatbb/corrections/qa08401.asp?qaDate=2006

*Piercy, F., & Lee, R. (1976). Effects of a dual treatment approach on the rehabilitation of habitual juvenile delinquents. *Rehabilitation Counseling Bulletin*, 19, 482–491.

Puzzanchera, C. (2009, December). *Juvenile arrests 2008.* : Office of Juvenile Justice and Delinquency Prevention Juvenile Justice Bulletin Accessed August 10, 2011 at http://www.ojjdp.gov/ojstatbb/default.asp

Quinn, M. M., Rutherford, R. B., Leone, P. E., Osher, D. M., & Poirier, J. M. (2005). Youth with disabilities in juvenile corrections: A national survey. *Exceptional Children*, 71, 339–345 2005-02490-008.

Roid, G. H., Ledbetter, M. F., & Messer, M. (2006). *Wide range achievement test, fourth edition — Progress monitoring version.* Lutz, FL: Psychological Assessment Resources, Inc.

Schrank, F. A., Mather, N., & Woodcock, R. (2004). *Woodcock–Johnson III(r) diagnostic reading battery.* : Riverside Publishing.

*Shoji, A. (2005). Comparative effectiveness of a selected juvenile delinquency prevention program across geographic locations. Doctoral Dissertation. Retrieved from ProQuest Dissertation and Theses Database. (UMI Number: 3180870).

Simpson, S. B., Swanson, J. M., & Kunkel, K. (1992). The impact of an intensive multisensory reading program on a population of learning-disabled delinquents. *Annals of Dyslexia*, 42(1), 54–66.

Sugai, G., & Horner, R. H. (1999). Discipline and behavioral support: Practices, pitfalls, and promises. *Effective School Practices*, 17(4), 10–22.

Vacca, J. S. (2008). Crime can be prevented if schools teach juvenile offenders to read. *Children and Youth Services Review*, 30, 1055–1062. http://dx.doi.org/10.1016/j.childyouth.2008.01.013doi:10.1016/j.childyouth.2008.01.013.

*Valentine, J., Griffith, J., Ruthazer, R., Gottlieb, B., & Keel, S. (1998). Strengthening causal inference in adolescent drug prevention studies: Methods and findings from a controlled study of the Urban Youth Connection Program. *Drugs & Society*, 12(1–2), 127–145.

Vaughn, S., Denton, C. A., & Fletcher, J. M. (2010). Why intensive interventions are necessary for students with severe reading difficulties. *Psychology in the Schools*, 47(5), 432–444.

Wanzek, J., Wexler, J., Vaughn, S., & Ciullo, S. (2010). Reading interventions for struggling readers in the upper elementary grades: A synthesis of 20 years of research. *Reading and Writing*, 23(8), 889–912. http://dx.doi.org/10.1007/s11145-009-9179-5.

*Weeks, S. Z. (1984). The effects of Sierra II, an adventure probation program, upon selected behavioral variables of adolescent juvenile delinquents. Unpublished Doctoral Dissertation. University of Virginia, Charlottesville, Virginia.

Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: A meta-analysis. *Psychological Bulletin*, 132(1), 132–149.

*White, Jerry D. (1979). Recidivism and institutional adjustment of institutionalized male juvenile delinquents involved in a vocational training program. Unpublished Doctoral Dissertation. North Texas State University, Denton, Texas.

Wiederholt, J., & Bryant, B. (2001). *Gray oral reading tests* (Fourth Edition). Austin, TX: PRO-ED Publishers.

Wiesner, M., Kim, H., & Capaldi, D. (2005). Developmental trajectories of offending: Validation and prediction to young adult alcohol use, drug use, and depressive symptoms. *Development and Psychopathology*, 17(1), 251–270. http://dx.doi.org/10.1017/S0954579405050133.

Wilson, D. B., Gottfredson, D. C., & Najaka, S. S. (2001). School-based prevention of problem behaviors: A meta-analysis. *Journal of Quantitative Criminology*, 17(3), 247–272. http://dx.doi.org/10.1023/A:1011050217296.

Wilson, S., & Lipsey, M. W. (2001). The role of method in treatment effectiveness research: Evidence from meta-analysis. *Psychological Methods*, 6(4), 413–429.

Wilson, S., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine*, 33(Suppl. 2), S130–S143. http://dx.doi.org/10.1016/j.amepre.2007.04.011.

Wilson, S., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive behavior: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 71(1), 136–149. http://dx.doi.org/10.1037/0022-006X.71.1.136.

Wilson, S., Lipsey, M. W., & Soydan, H. (2003). Are mainstream programs for juvenile delinquency less effective with minority youth than majority youth? A meta-analysis of outcomes research. *Research on Social Work Practice*, 13(1), 3–26.

Zabel, R. H., & Nigro, F. A. (1999). Juvenile offenders with behavioral disorders, learning disabilities, and no disabilities: Self-reports of personal, family, and school characteristics. *Behavioral Disorders*, 25(1), 22–40.

Zabel, R. H., & Nigro, F. A. (2007). Occupational interests and aptitudes of juvenile offenders: Influence of special education experience and gender. *Journal of Correctional Education*, 58, 337–356.