EDUCATION FOR INCARCERATED JUVENILES: A META-ANALYSIS

Jennifer L. Steele, American University
Robert Bozick, RAND Corporation
Lois M. Davis, RAND Corporation
2015

Abstract

Based on screenings of 1,150 manuscripts, we synthesize evidence from 18 eligible studies of educational interventions implemented within juvenile correctional facilities. The studies include five intervention categories: remedial academic instruction, computer-assisted instruction, personalized academic instruction, vocational education, and GED completion. Effectiveness is measured in terms of four outcomes: academic performance in reading or mathematics, diploma completion, post-release employment, and post-release recidivism. Focusing on studies with the strongest basis for causal inference, we find positive and statistically significant effects for computer-assisted instruction in raising reading comprehension, and for personalized learning in improving diploma completion and post-release employment. These findings are driven by large and well-executed randomized trials of Scholastic's Read 180 curriculum and Florida's Avon Park Youth Academy. Despite the limited research base, these studies suggest that it is possible to undertake rigorous research in juvenile facilities about programs that best improve the outcomes of young offenders.

Introduction

Insofar as federal protection exists for access to a public education in the United States, it applies to youth involved in the criminal justice system (Lhamon, 2014; Lhamon & Gupta, 2014). Because youth typically cannot attend their local schools while incarcerated, correctional facilities must provide an alternative educational system to serve them. This is challenging because youth are transient within the justice system, and because they arrive with widely varying academic and emotional needs (Sedlak & McPherson, 2010).

In light of that challenge, this article aims to help policymakers and practitioners understand the evidence about what works in correctional education for juveniles. Evidence is drawn from comparison-group and single-case design studies of academic or vocational-education program effectiveness for incarcerated youth. Using a systematic review procedure that included 1,150 title-and-abstract screenings and136 full-text screenings, we synthesize findings from 18 studies published between 1980 and 2011 that meet evidence standards. Effectiveness is defined in terms of four outcomes relevant to correctional educators: test scores in reading and mathematics, diploma completion, post-release employment, and post-release recidivism. The interventions considered fall into five categories: remedial academic instruction, computer-assisted instruction, personalized academic instruction, vocational education, and GED completion. Given the small number of eligible studies conducted within juvenile

correctional facilities, we briefly contextualize our synthesis with a discussion of the broader evidence in each intervention category. We conclude with implications for both research and policy.

Overview of Juvenile Correctional Education

In 2011, about 61,000 individuals below age 21 were incarcerated in juvenile facilities on any given day in the United States. This figure represents roughly 0.25 % of the national population age 15–20 (U. S. Census Bureau, 2011; Sickmund et al., 2013).[1] The rate of youth incarceration in the United States is more than three times the highest rates in other developed nations (Hazel, 2008), but it is much lower than the U.S. adult incarceration rate, which was about 1 % in 2011 (Glaze & Parks, 2012, U.S. Census Bureau, 2011). Nevertheless, the number of youth in juvenile facilities in the United States has declined remarkably in the past 17 years, dropping from about 105,000 in 1997 (Sickmund et al., 2013). The decline may partially reflect growing evidence that incarceration exacerbates recidivism risk in young people (Aizer & Doyle, 2013).

In this article, we define *incarcerated youth* as individuals under age 21 who are legally assigned to correctional facilities as a result of arrest, detainment for

[1] In addition, about 2,437 youth under 18 were held in adult facilities at the midpoint of the closest available year, 2010. The number of individuals aged 18-20 in adult facilities is more difficult to ascertain, but extrapolating from the 2015 share of inmates aged 18-20 in federal prisons (1.2%) (Federal Bureau of Prisons, 2015), we might approximate that about 19,000 individuals aged 18-20 were held in adult facilities at the midpoint of 2010 (Guerino et al., 2011).

court proceedings, adjudication by a juvenile court, or conviction in an adult criminal court (Office of Juvenile Justice and Delinquency Prevention, 2013).

Among individuals held in juvenile correctional facilities in the United States in 2011, about 86 %were male, and youth of color were markedly overrepresented. Forty percent of incarcerated youth were black and 23 %were Hispanic, as compared with about 13 percent and 17 percent, respectively, in the U.S. population at large (Sickmund et al., 2013; U.S. Census Bureau, 2013).[2] About 30 %of youth incarcerated in juvenile correctional facilities in 2011 were under the age of 16. Another 55 % were ages 16 or 17, and 14 % were ages 18 to 20 (Sickmund et al., 2013).

Juvenile offenders hail disproportionately from challenging circumstances (Sedlak & McPherson, 2010), and are much more likely than their nonoffender counterparts to have emotional problems (Snyder & Sickmund, 2006) and substance-abuse histories (Sedlak & McPherson, 2010). These problems are compounded by offenders' comparatively weak academic skills. The average reading ability of incarcerated youth has been estimated at the fourth-grade level, placing them 5 years behind average grade-level targets (Project READ, 1978). Incarcerated youth are also more likely than their counterparts to be learning disabled; one synthesis suggested that between 30 and 50 % of incarcerated youth

[2] Figures for incarcerated youth are based on 2011 data; comparison data for the U.S. population come from 2012.

have special education needs, as compared with approximately 10 % of nonincarcerated youth (Mears and Aron, 2003). Given the disproportionate representation of students with disabilities in juvenile correctional facilities and the high cost of educating these students (Chambers et al., 2004), juvenile correctional facilities often struggle to adequately serve the special needs of their students (Leone, 1994; Pasternak et al., 1988).

Nationally, the longterm trajectories of juvenile offenders are murky because states track juvenile recidivism using different metrics and different subsets of offenders, and some states do not make such data available at all. According to a 2006 report that used data from Florida, New York, and Virginia, the 12-month rearrest rate among released juvenile offenders in either the juvenile or adult system was 55 %. Using data from eight states, the same report estimated that 33 % of juvenile offenders were readjudicated in a juvenile court or reconvicted in an adult criminal court within 12 months after release from a juvenile facility (Snyder & Sickmund, 2006).

Other Reviews of Interventions for Juvenile Offenders

In this article, we synthesize research evidence about how best to educate youth held in correctional facilities. Focusing strictly on juvenile correctional education interventions, our review is designed to complement those that have preceded it. In 2009, Lipsey published a meta-analysis of interventions designed to reduce juvenile delinquency. His study, which was based on 548 effect

estimates, included a wide array of interventions (not just educational) and was not limited to studies conducted within correctional facilities. He also limited his analysis to studies that focused on recidivism as the dependent variable of interest, finding that programs with therapeutic rather than coercive approaches were more effective in reducing recidivism, as were programs that served higher risk youth, and programs with higher quality implementation. Later, Sander and colleagues (2012) examined the effects of a broad array of interventions for juvenile offenders, not limited to education, and including studies conducted within and outside of correctional facilities. Their review, which included 134 effect estimates, diverged from Lipsey's by focusing on academic rather than recidivism outcomes. Noting the limited number and quality of eligible studies, it did not find statistically significant program effects on achievement, attendance, or attitudes. Our synthesis differs from both in that it is limited to studies of academic and vocational education interventions, and these interventions must have been implemented within correctional facilities.

Building on a small review of reading interventions in correctional settings by Krezmien & Mulcahy (2008), Wexler et al. (2014) synthesized 16 studies that focused on academic interventions in juvenile correctional facilities and looked solely at academic outcomes. Though our review partially overlaps with that of Wexler et al. (2014) in terms of the manuscripts considered, it is more inclusive in three key ways. First, we include both academic and vocational interventions

rather than academic interventions only. Second, we consider not only academic achievement outcomes but also outcomes related to employment and recidivism. Third, the Wexler et al. (2014) review focused only on studies published in peer-reviewed journals. Because intervention studies are often published in research reports rather than journals, and because we are concerned about minimizing publication bias (Borenstein et al., 2009), we include any study that meets our methodological criteria, regardless of whether it is published in a peer-reviewed journal. With these broader criteria, we include 9 studies that were not part of the Wexler et al. synthesis.

We are also more restrictive than the Wexler et al. (2014) review in one key way. Our evidence standards consider the study's basis for causal inference, using standards adapted from the Maryland Scientific Methods Scale (Sherman et al., 1997) and the U.S. Department of Education's What Works Clearinghouse (WWC) standards (What Works Clearinghouse, 2014), as described in the methods section below. Because we exclude studies that lack a comparison group unless they meet WWC standards for single-case designs, we exclude seven of the 16 studies that Wexler et al. include, though we do footnote the reasons for their exclusion in the relevant sections below.

Study Sample and Methods

Document Identification

Our search for extant research included queries of several databases, including Education Resources Information Center (ERIC), Education Abstracts, Criminal Justice Abstracts, National Criminal Justice Reference Service Abstracts, Academic Search Elite, EconLit, Sociological Abstracts, and Google Scholar. Our search was limited to studies conducted in the United States and released from 1980 through 2011; it required at least one descriptor from each of the following sets of terms: (Set 1) youth or juvenile; (Set 2) juvenile justice, prison, jail, incarcerat* (where the asterisk allows for different word endings), detention center, or corrections; and (Set 3) education, academic, diploma, GED, literacy, math, reading, science, job skills, job training, apprentice*, vocational education, voc tech, occupational education, career and technical education, workforce (or work force) development, workforce training, workforce preparation, or school to work.

We supplemented the manuscripts identified through these searches with manuscripts cited by existing literature reviews on the topic of juvenile correctional education. Altogether, the document search resulted in 1,150 citations for title and abstract screening, as shown in Figure 1, which summarizes our search and screening process.

<Insert Figure 1 about here>

Eligibility Assessment

Eligible manuscripts had to describe the effects of an academic or vocational intervention on incarcerated juveniles, where the definition of juveniles was permitted to be defined by the manuscripts themselves or to include participants under age 21. The manuscripts were also required to be primary, empirical studies rather than literature reviews or opinion pieces. Each manuscript was screened for these criteria independently by two doctoral students using the research synthesis software package DistillerSR. Prior to screening, screeners were trained by a senior team member using sample articles. Interrater reliability during this training period rose from 63% to 92%. All conflicts between screeners were resolved in DistillerSR by a senior team member who served as a moderator. The screening process yielded 157 manuscripts eligible for full-text screening. Of these, 21 were duplicates or had full texts that could not be located, resulting in 136 manuscripts that underwent full-text screening.

To pass full-text screening and be deemed eligible for inclusion in the meta-analysis, the study was required to: (1) evaluate an eligible intervention within an eligible population and setting; (2) measure success of the program using an eligible outcome measure; and (3) employ an eligible research design. An *eligible intervention* was defined as any academic or vocational education intervention program. An *eligible population* was defined as consisting primarily of individuals below age 21. An *eligible setting* was any facility, regardless of

jurisdiction (state, local, etc.), to which juveniles were confined due to arrest, court proceedings, or adjudication or conviction. Eligible interventions were limited to academic or vocational education programs. Eligible interventions were permitted to include an aftercare (i.e., postrelease) component, but the interventions had to be delivered primarily within the correctional facility. Interventions that did not provide instruction in academic or vocational skills, such as mentoring programs, substance abuse programs, and mental health programs, were excluded.

We defined *eligible outcome measures* as any measure of recidivism (e.g., rearrest, reconviction, or reincarceration), postrelease employment, academic attainment (e.g., GED or high school completion), or academic performance in reading or mathematics (e.g., test scores).

We included two types of studies in the definition of *eligible research design*. The first and most common type was a *comparison-group design* in which a group of incarcerated juveniles who received an intervention were compared with a group of incarcerated juveniles who did not, or who received a different version of the intervention. The second type was a *single-case design*, which is discussed later in this section.

For comparison-group designs, we rated the rigor of the studies using two scales that closely correspond to one another—the Maryland Scientific Methods Scale, which is based on a 1997 University of Maryland report to Congress about

what works in crime prevention (Sherman et al., 1997), and the evidence rating scale used by the U.S. Department of Education's What Works Clearinghouse (2014). On both, assessments of rigor reflect the extent to which the designs protect against unobserved differences between the treatment and comparison groups that are correlated with the outcome of interest (i.e., selection bias). Table 1 summarizes the standards employed for both scales. Randomized trials with low attrition constitute the most rigorous of these types of designs, because randomizing the two groups renders the treatment and comparison group alike in expectation (Shadish, Cook, & Campbell, 2002). We assigned these designs a 5 (the highest rating) on the Maryland Scale, and a *meets standards* rating on the WWC scale. Studies that demonstrate very close matches between treatment and comparison groups on relevant observable characteristics (at minimum, age, prior offenses, baseline education level, and time to data collection) are awarded a 4 on the Maryland Scale and a *meets standards with reservations* rating on the WWC scale. Studies that do not demonstrate strong baseline matches (within a 20th of a standard deviation for the aforementioned variables) but that attempt to control for observed baseline differences earn a 3 on the Maryland Scale, but *do not meet standards* on the WWC scale. Studies that do not attempt to control for observed baseline differences between the treatment and comparison groups earn a 2 on the Maryland Scale, and do not meet WWC standards. The Maryland Scale assigns a rating of 1 to studies that do not include a comparison group because they include

no way to estimate what would have happened to the treatment group in the absence of treatment, and these studies are not eligible for WWC review. We exclude Maryland Level-1 studies from our review.

<Insert Table 1 about here>

We make one notable exception to the comparison-group requirement, and that is for studies that use a class of approaches called *single-case designs*. Single-case designs are commonly employed in special-education research, where large samples are often unavailable for intervention evaluation (Kratochwill et al., 2010). They involve systematically introducing an intervention with one or a few students in an effort to demonstrate causal effects. These studies include a large number of pre- and postintervention outcome measurements, allowing students to function as their own controls, similar to student fixed-effect models in econometric research (Wooldridge, 2002). Focusing on one or a handful of participants, these designs typically lack statistical power for conventional hypothesis testing. However, insofar as researchers can establish a clear preintervention trend, then deviations from that trend in the presence of the intervention can be causally attributed to the intervention itself. The U.S. Department of Education's What Works Clearinghouse (WWC) has therefore established specific standards for the rigor of single-case design studies (Kratochwill et al., 2010). We follow these standards when rating the single-case designs included in our analysis. This means that we assign a Level 5 rating on

the Maryland Scale to studies that receive the highest rating (*meets standards*) under WWC standards for single-case designs, because these studies demonstrate a strong basis for inferring that observed effects are causal.

Scientific Review and Synthesis

Only 18 of the 136 studies subjected to full-text screening were deemed eligible for inclusion in the research synthesis. Eligible studies were reviewed independently by two Ph.D.-level experts in juvenile correctional education who participated in a 2-day training on the data extraction protocol (available as Appendix C in Davis et al., 2014). These data were then integrated and cleaned, with reference to the original texts, by a doctoral student. Next, the integrated and cleaned data were reviewed by a senior member of the research team, and extracted data from each study were closely checked against the original articles. Based on the extracted data, each study was rated for rigor on the Maryland Scale and WWC scale. Data from the 18 eligible studies were organized and summarized by intervention type; summaries of each are shown in Appendix Table A1. This table includes information about each study, including descriptions of the treatment and comparison conditions, the demographics of the study population, the size of the treatment and comparison groups, a brief description of intervention duration and frequency, summaries of the reported effect estimates, and indications of the Maryland Scale rating we assigned to each study.

Analytic Approach

A key challenge facing this research synthesis is that the number of eligible studies is small, and the interventions under consideration are heterogeneous. In light of these constraints, the use of a quantitative meta-analytic approach is not ideal. However, we have provided a qualitative review of the eligible interventions as chapter 3 in Davis et al. (2014). Here, our aim is to summarize these findings more succinctly. To that end, we report on mean estimates overall and for each intervention category using a quantitative meta-analytic approach, while still noting the ways in which their generalizability is limited.

We deal in part with the heterogeneity of interventions by breaking them into five substantive categories, but within any given category, the number of studies is especially small. Moreover, each study considers only a subset of the dependent variables of interest, so not all dependent variables are represented within each intervention category, and in some cases, a particular dependent variable is examined by only one study in a given category. The consequence is that our ability to generalize beyond the studies is limited, and the fact that so few studies use designs that warrant causal inference exacerbates this limitation. For this reason, this article aims not only to report on average effects, but also to describe the attributes of interventions for which evidence is particularly strong.

To summarize our findings for each intervention category and available outcome, we aggregate estimates using a DerSimonian-Laird random effects

model, in which each estimate is weighted by the inverse of its variance and by its difference from the mean effect (DerSimonian & Laird, 1986).[3] Cochrane's Q is used as a test of effect heterogeneity across studies, where the null hypothesis is that the true effects are homogeneous (DerSimonian & Laird, 1986). With such a small sample of studies, however, we may lack power to reject the null hypothesis of study homogeneity even if true effects are heterogeneous. In this case, we anticipate heterogeneous effects because of the diversity in interventions and methodological rigor.

We generate meta-analytic estimates using the metaan package for Stata (Kontopantelis & Reeves, 2010). Two of the studies provide multiple estimates of reading effects. To address the dependence of same-study estimates within the same outcome category, we calculate a study-average effect and standard error following Borenstein and colleagues (2009), and assuming a correlation of 0.7 between measures of different reading skills (Fuchs et al., 2001). Also, in our discussion of estimates by intervention category, we actually examine two types of reading outcomes separately: *word-level reading skills* such as sight-word

[3] The variance is the inverse of the squared standard error. The standard error is calculated as $\sqrt{\frac{1}{p_t n_t} + \frac{1}{(1-p_t)n_t} + \frac{1}{p_c n_c} + \frac{1}{(1-p_c)n_c}}$ for dichotomous outcomes, with p_t indicating treatment group probabilities and n_t indicating treatment group sample sizes, etc. For continuous outcomes, it is calculated as $\sqrt{\frac{n_t + n_c}{n_t n_c} + \frac{d^2}{2(n_t + n_c)}}$, where d^2 represents the squared standardized mean difference (Bland & Altman, 2000; Wilson, 2011).

identification and decoding (sounding out unfamiliar words), and *reading comprehension skills*, such as understanding or drawing inferences about the main ideas of a passage. We take this approach because some of the reading interventions under consideration concentrate primarily on one type of skill or the other.

In the next sections, we first present overall impact estimates across all study categories for the dependent variables of interest: intervention effects on academic test scores, on diploma completion rates while incarcerated, on employment rates after release, and on recidivism rates after release. Academic test score effects are scaled in standard deviation units,[4] and effects on the other three dependent variables are presented as odds ratios. The *odds of an event occurrence* are defined as the probability that it occurs divided by the probability that it does not, and the *odds ratio* is defined as the odds of the event for the intervention group divided by the odds for the comparison group. An odds ratio greater than one indicates higher odds of the event (e.g., diploma completion, employment, or recidivism) among the intervention group, and an odds ratio between 0 and 1 indicates that the odds of the event are lower among the intervention group. An odds ratio of 1 indicates no effect. We also consider the extent to which the

[4] Kratochwill and colleagues (2010) acknowledge the difficulty of presenting summary statistics for single-case design studies. In order to include these studies in our meta-analytic estimates, we calculate standardized mean differences using the preintervention and postintervention means of each student, and we calculate standard errors based on the total number of students in the study.

overall estimates depend on the methodological rigor of the studies (based on Maryland Scale ratings). We use Egger and Begg tests to assess evidence of publication bias. An Egger test is the intercept in a regression of the effect size divided by its standard error on the inverse of its variance. A Begg test, which we use in categories with only two studies, is the rank correlation between effect sizes and their variances (Egger et al., 1997).

Results

Before presenting estimates for each intervention category, we first briefly consider overall meta-analytic estimates across the 18 eligible studies for the four dependent variables of interest: academic performance, high school diploma or GED completion, postrelease employment, and postrelease recidivism. Given that these studies focus on distinct types of interventions, the overall estimates can be interpreted as the average effects of the juvenile correctional interventions in the eligible studies relative to the default educational programs in the study settings. Across all of the studies deemed eligible for synthesis, we have 14 reading or math effect estimates from 10 studies, two diploma completion estimates from two studies, two employment estimates from two studies, and eight recidivism estimates from eight studies. Forest plots of the estimates and confidence intervals for each study are shown in Appendix Figure A1, along with the weighted meta-analytic means. The size of the squares represents the weight of each estimate in the analysis, with more precise estimates receiving greater weight.

Table 2 presents meta-analytic estimates and confidence intervals for each outcome, as well as the Cochrane's Q test statistic (and an associated p-value) for effect heterogeneity. We are able to reject the null hypothesis of homogeneous effects only for recidivism (among all 8 studies), but as noted above, the small samples restrict the power of the heterogeneity tests. We continue to employ a random-effects approach throughout, in which we conservatively assume that the true effects are heterogeneous.

<Insert Table 2 about here>

In Table 2, we find that among the 14 reading or math effect estimates eligible for inclusion, the weighted mean effect is approximately 21 % of a standard deviation. All of these estimates are from studies with a Maryland Scale level of 3 or higher (meaning they at least employ statistical controls for observed baseline differences), but when we limit the analysis just to studies with a Maryland Scale Level 5 (low-attrition randomized trials or strong single-case designs), we obtain the same estimate of about 0.21. Both estimates are statistically significant at the 5 % level, with a confidence interval for the highest-rigor studies of 0.13 to 0.29 of a standard deviation.

We have only two studies that use diploma completion as a dependent variable, and two that use employment as a dependent variable. The estimated odds of diploma completion associated with participation in the intervention programs are more than three times the odds for those participating in default

programs, and the estimated odds of post-release employment for the intervention programs are 1.4 times the odds for those in the default programs. For both outcomes, estimates appear similar for the higher and lower rigor studies, but these estimates become useful only when we explore the program attributes in the next section.

Among the eight studies examining post-release recidivism effects, the estimated odds of recidivism for students participating in the programs of interest are only 70 % of the odds of those participating in default programs, but this effect is not statistically distinguishable from zero, and moreover, five of the estimates come from studies with a Maryland Scale rating of only 2. Limiting the analysis just to studies with a rating of 3 or 5, the estimate is a nonsignificant 0.95, and it is 1.04 (also nonsignificant) for the Level 5 study alone.

Table 2 also includes test statistics for publication bias. If publication bias leads smaller studies to be published only when they show desirable effects, then the estimates' precision and magnitude may be negatively correlated (or positively correlated in the case of recidivism estimates). If publication bias is not present, there should be no relationship between precision and magnitude, meaning that our Egger and Begg tests for such bias would not reject the null hypothesis. Though none of the publication bias tests are statistically significant in Table 2, we have very few studies in each category upon which to base a hypothesis test. We also include funnel plots in Appendix Figure A2, in which we

look for asymmetry in the relationship between precision and magnitude. Less vertical symmetry at the bottom of the plots than at the top indicates possible publication bias. In this case, the academic outcomes in the top left panel look reasonably consistent, and it is difficult to make a determination about diploma completion and employment with only two studies per plot. However, there is some visual indication of publication bias in the recidivism estimates, even though the Egger test coefficient is not statistically significant. This is consistent with the weaker general rigor of the recidivism studies in our sample relative to the studies with academic outcomes.

We turn now to a discussion of findings for each intervention category:remedial academic instruction, computer-assisted instruction, personalized academic instruction, vocational training, and earning a GED while incarcerated. These findings are summarized in Table 3, which presents weighted meta-analytic means, confidence intervals, and Cochrane's Q statistics for each dependent variable examined within each intervention category. Table 3 also notes the number of effects and studies underlying each estimate, and the Maryland Scale ratings of the studies. For simplification, we classify the basis for causal inference for a given intervention category and outcome as *high* if some underlying studies rate a 5 and none rates lower than a 3; as *moderate* if all rate a 3, or some rate a 5 and some rate a 2; as *lower* if some rate a 3 and some a 2; and as *lowest* if all rate a 2. (In the sample, no studies used matched quasi-

experimental designs that would have earned a rating of 4.) We consider the evidence for a category to be both compelling and significant if it is based on at least one Level 5 study and no Level 2 studies, and if it reaches statistical significance at the 5 % level. Because we have so few studies from which to extrapolate in each intervention category, we introduce each category with a brief discussion of relevant research conducted outside of juvenile correctional settings.

<Insert Table 3 about here>

Remedial Academic Instruction

In recent years, a growing body of literature has questioned the effectiveness of remedial education for improving student outcomes. But this literature has largely focused on postsecondary education, where remedial education can slow a student's progress and increase the cost of earning a degree, thereby acting as a potential deterrent to degree completion (Caldagno & Long, 2008; Martorell & McFarlin, 2011; Scott-Clayton & Rodriguez, 2012). In secondary education, the need to remediate the learning gaps of students who fall behind seems less controversial, the question being how best to do so.

The three remedial academic interventions eligible for inclusion in our analysis are Corrective Reading in three studies (Allen-DeBoer et al., 2006; Drakeford, 2000; Scarlato & Asahara, 2004), the Orton-Gillingham structured remedial reading program in one study (Simpson et al., 1992), and a program classified as "remedial education," without descriptive details (Archwamety &

Katsiyannis, 2000) in a fifth study. Corrective Reading is a commercially available, intensive reading program that emphasizes direct instruction and is designed for students whose reading skills are below grade level (McGraw Hill Education, 2013). In 2007, based on studies outside of correctional settings, the WWC deemed the curriculum to have potentially positive effects on alphabetics (e.g., phonics and decoding) and fluency (e.g., rate and accuracy), but no discernible effects on comprehension (What Works Clearinghouse, 2007, drawing on Torgesen et al., 2006). The second remedial intervention is the Orton-Gillingham curriculum, a commercial reading program that targets students with dyslexia. Two research syntheses on the Orton-Gillingham conducted mainly outside of correctional settings found mixed effects (Ritchey & Goeke, 2006) or insufficient rigor on which to base conclusions (What Works Clearinghouse, 2010).

In our study sample, we estimate that the weighted mean effect of these remedial programs on *word-level reading skills*, such as correctly identifying familiar words or sounding out unfamiliar words, is more than half a standard deviation, at 0.534. Though this estimate's basis for causal inference is reasonably high, with two Level 5 and one Level 3 studies, the combined sample size across the three studies is only 19, and the estimate is not statistically distinguishable from zero. The estimated effect of remedial education on *reading comprehension* is also large, at 0.526 of a standard deviation, and in this case *is* statistically

distinguishable from zero at the 5 & level. However, the basis for causal inference is only moderate, with two Level 3 studies, and with a combined sample of 72 participants, the confidence interval is quite wide, ranging from about 6 to 100 % of a standard deviation. Finally, among the two studies that examine the association between remedial education and *recidivism* (combined n=568), we estimate a statistically nonsignificant odds ratio of 0.936. An important caveat, however, is that this estimate has a lower basis for causal inference, emerging from a large Level 2 and small Level 3 study.

Given that the only statistically significant estimate in this category—on reading comprehension—has just a moderate basis for causal inference, we do not find compelling or statistically significant evidence in support of the remedial interventions in the analysis.

Computer-Assisted Instruction

Meta-analyses of the effects of computer-assisted learning outside of correctional-education settings have yielded mixed indications of these programs' effectiveness in raising student achievement (Soe, Koki, & Chang, 2000; Slavin et al., 2008). But computer-based curricula vary widely, making it difficult to generalize about them as a class of interventions. We therefore focus our discussion on the three computer-assisted interventions that are included in our analysis: Read 180 (Loadman et al., 2011), Fast ForWord (Shippen et al., 2012),

and Tune in to Reading (Calderone et al., 2009). All three are rated Level 5 on the Maryland Scale, meaning that our basis for causal inference is strong.

Read 180, published by Scholastic (n.d.), is a complete reading curriculum for upper elementary through high school that includes an adaptive, computer-assisted component as well as teacher-led instruction, independent reading, and small-group reading. In previous reviews, the What Works Clearinghouse (2009) found potentially positive effects of Read 180 on reading comprehension and general literacy achievement, and Slavin et al. (2008) estimated a weighted mean effect of 0.24 of a standard deviation, though none of the underlying studies in either review was conducted in a juvenile correctional setting. .

The Fast ForWord software intervention is published by Scientific Learning Corporation. Unlike Read 180, it is a completely computer-based curriculum designed for beginning readers (Scientific Learning Corporation, 2004). The developer of Fast ForWord, Scientific Learning Corporation (2004), did study the intervention with 29 youth incarcerated by the Virginia Department of Correctional Education, finding reading gains of 1.5 grade equivalents over a four-to-ten month period, but in the absence of a comparison group, the study rates a 1 on the Maryland Scale and is not eligible for inclusion in our analysis. Reviewing studies conducted outside of juvenile correctional settings, the What Works Clearinghouse (2013) found positive effects on alphabetics (e.g.,

decoding), no effect on reading fluency (e.g., rate and smoothness), and mixed effects on reading comprehension.

The final eligible computer-assisted intervention in our analysis is Tune in to Reading (TIR). Published by Electronic Learning Products, the software promotes fluency by teaching students to sing written words with the correct pitch and tone (Calderone et al., 2009). A small matching study without random assignment found positive effects as large as 0.98 of a standard deviation, though the study was not carried out in a correctional-education setting (Biggs et al., 2008).

As shown in Table 3, our synthesis of computer-assisted instruction studies conducted in correctional settings includes one estimate for word-level reading skills, based on the FastForWord randomized trial, which found a negative and non-statistically significant effect of -0.123 of a standard deviation. Focusing on reading comprehension estimates across all three studies, we find a positive and statistically significant effect of 0.2 of a standard deviation. This effect is driven mainly by the Read 180 study, which included 1,245 participants, as compared to only 51 participants in the Fast ForWord study, and 103 participants in the Tune in to Reading study. Still, effect sizes were nearly identical, at 0.21 of a standard deviation, for Read 180 and Tune in to Reading. In contrast, Fast ForWord had a negative reading comprehension estimate of -0.17.

Given that all eligible studies in this category come from well executed randomized trials (Level 5 studies), and that the weighted mean effect on reading

comprehension is statistically significant and positive, we conclude that computer-assisted instruction (especially as represented by Read 180 and Tune into Reading) holds promise for improving the reading-comprehension skills of incarcerated juveniles.

Personalized Academic Instruction

Personalized learning is a broad term in education literature, indicating that instruction is adjusted to fit the unique needs and developmental trajectories of each student. The notion of personalized learning as especially effective is based on Bloom's (1984) "two-sigma problem," in which he argued that individualized tutoring yielded gains two standard deviations higher than traditional classroom instruction. A subsequent meta-analysis placed the estimate closer to 0.79 of a standard deviation (VanLehn, 2011), but this is still a large effect. The challenge is that one-to-one teacher-student ratios are costly, so the question of how to cost-effectively personalize learning remains open.

Our meta-analysis includes five studies of personalized learning in juvenile correctional facilities, in which instruction was tailored to the individual needs of each student. The studies include a large randomized trial evaluating the Avon Park Youth Academy (National Council on Crime and Delinquency [NCCD], 2009). This academy, operated by the Florida Department of Juvenile Justice, used an intensive, personalized instructional model tailored to each student's academic development. It also included vocational programming within the

facility and aftercare upon release, including a 76 % reduction in probation officers' caseload. The study, which included 714 participants, had no attrition and rates a 5 on the Maryland Scale. The category also includes a small study of the Regional Youth Educational Facility (RYEF) in San Bernardino County, California, a program that included six months of intensive, personalized instruction within the juvenile facility, followed by 4 to 6 months of aftercare supervision by a probation officer familiar with the youth through RYEF (Skonovd et al., 1991). This study, which included 45 participants, did not use random assignment or statistical controls, so rates only a 2 on the Maryland Scale. Two other eligible personalized education programs include a 143-student study of instruction customized to students' rates of progress (Mayer & Hoffman, 1982), and a 38-student study of individualized peer-tutoring program with a 1:1 or 1:2 ratio relative to teacher-managed instruction with a ratio ranging from 1:3 to 1:7 (Kane & Alley, 1980). Neither of these studies used random assignment, but both included pretest adjustments, resulting in a Level 3 Maryland Scale rating. Finally, the category includes a 20-student randomized trial of Corrective Reading in which a student-teacher treatment ratio of 4:1 was compared to a less-personalized 12:1 ratio in the comparison group (Houchins, 2008). This study has low enough attrition to rate a 5 on the Maryland Scale, and includes pretest score adjustments.

In Table 3, we find that only the estimates for diploma completion and employment are statistically significant, with odds ratio estimates for personalized learning relative to default programs of 3.42 for diploma completion and of 1.45 for postrelease employment. These are substantial effects based on a well executed random-assignment research design (NCCD, 2009), though they are admittedly based on only a single evaluation, the Avon Park program in Florida. In contrast, the estimated positive effect of 0.18 of a standard deviation for academic learning in reading or math, and the estimated reduction in recidivism by 42 % (given the odds ratio of 0.58), are not statistically distinguishable from zero. The latter is driven by the Level 2 RYEF study in San Bernardino County. The Level 5 Avon Park study found recidivism rates that were slightly higher (though not statistically significantly) among those randomized to the program, despite their having notably higher diploma completion and postrelease employment rates.

Considering the category as a whole, we conclude that there is compelling and statistically significant evidence in support of personalized instruction for improving diploma completion and postrelease employment, but this evidence pertains only to the Avon Park Youth Academy, which was a particularly multifaceted and intensive program.

Vocational /Career and Technical Education

Among the population of U.S. secondary school students at large, the prevalence of vocational training, now commonly termed *career and technical education* (CTE), declined between 1982 and 2004, with vocational credits accounting for 21 % of the credits earned by high school graduates in 1982, versus only 14 % in 2004 (U.S. Department of Education, 2013). (We use the term *vocational education or CTE* to reflect the language in the studies we synthesize in this section, though CTE is the more contemporary term.) In part, the decline was a response to concerns that lower achieving students were being tracked into vocational pathways that did not prepare them to succeed in an increasingly competitive and dynamic labor market (U.S. Department of Education, 2013). Evidence on the effectiveness of vocational education or CTE in raising academic outcomes is somewhat mixed (Bozick & Dalton, 2013; Kemple & Willner, 2008; Neild, Boccanfuso, & Byrnes, 2013). But insofar as a lack of marketable skills increases the appeal of criminal behavior (Becker, 1968), it is possible that juveniles involved in the criminal justice system may be especially likely to benefit from programs that emphasize vocational skills.

In his aforementioned meta-analysis of 548 juvenile crime-reduction effect estimates for juvenile offenders, Lipsey (2009) considered studies of a variety of programs (educational and otherwise) designed to reduce youth recidivism, 22 % of which were conducted in juvenile facilities. He found that *skill-building interventions*—defined to include behavior management, cognitive-behavioral

therapy, social skills training, challenge programs, academic training, and job-related interventions including vocational training—reduced subsequent recidivism by about 6 percentage points, though the effect was not statistically significant.

Our own analysis of vocational and CTE programs identified three eligible studies of participation in a stand-alone vocational education or CTE program in a juvenile correctional facility. In a large but observational (i.e., nonrandomized) study, Roos (2006) examined the employment and recidivism rates for participants of the Re-Integration of Offenders–Youth (RIO-Y) career development course operated by the Texas Youth Commission. The RIO-Y study is unique in our review in that the comparison group did not receive an alternative instructional program during the intervention period, which was possible because all were 18 years of age or older. The analysis adjusted for baseline demographic and risk-related covariates, so it warrants a Level 3 rating on the Maryland Scale. In another observational study, Wilson (1994) investigated the effects of vocational education or CTE in a juvenile correctional facility, but provided few details about the program. In addition, DelliCarpini (2010) capitalized on a policy shift at a New York State county jail, comparing youth exposed to a new set of vocational classes in business, drafting, and carpentry to those in earlier cohorts that had not had these classes. Neither the Wilson nor the DelliCarpini study adjusts for participants' baseline characteristics, so both rate a 2 on the Maryland

Scale. All three studies are substantially larger than previously discussed studies of academic interventions (ranging from 302 in the Wilson study to 1,502 in the Roos study), but they are generally less rigorous in terms of the basis for causal inference.

As shown in Table 3, the studies find positive and significant effects of vocational education for diploma completion, with an estimated odds ratio of 2.59, and for employment, with an estimated odds ratio of 1.38. Both estimates are statistically distinguishable from zero at the 5 % level, though their confidence intervals are wide, and both are based on a single study—a Level 2 in the case of diploma completion (DelliCarpini, 2010), and a Level 3 in the case of employment (Roos, 2006). Though the estimated odds of recidivism appear reduced, at a ratio of 0.67, this estimate is not statistically different from zero and has a fairly weak basis for causal inference.

Considering the vocational interventions in this analysis, we do not find compelling and significant evidence that these programs improved diploma completion or employment or reduced recidivism. Though the study estimates did show desirable effects for all three outcome types, the recidivism estimate is not statistically different from zero, and the research designs are not strong enough to warrant causal inferences.

GED Completion

Extant evidence on GED attainment suggests that, when separated from ability and motivation, earning a GED may increase the earnings of high school dropouts by 10-19 % (Tyler, Murnane, & Willet, 2000), but the question most pertinent to this analysis is whether there are benefits to earning a GED while incarcerated. In a rigorous study focusing on incarcerated *adults*, Tyler and Kling (2007), found that earning a GED in prison yielded earning gains of 15 % in the first 2 years after release, though the benefit dissipated after the second year. In addition, they found that most of the benefit came from participating in a GED education program rather than from actually earning the GED. In a meta-analysis also focusing on incarcerated adults, Davis et al. (2013) found that that participating in a high school diploma or GED program reduced the odds of recidivism by 30 %.

Two studies in our analysis examine the effects of GED attainment while incarcerated on postrelease recidivism. The particular challenge of these studies is that they focused not on participation in GED programs or on opportunities to earn a GED, but on actual attainment of the GED. This is a challenging independent variable because juveniles' completion of a GED while incarcerated may depend on many factors, including their length of stay, academic preparedness, and underlying motivation. The studies in this category are therefore especially vulnerable to selection bias because they use an independent

variable that is strongly linked to individual motivation and ability, and yet they are unable to adjust for these attributes.

The larger study, by Jeffords and McNitt (1993), examined reincarceration rates within 1 year after release among 1,717 youth in two juvenile correctional facilities in Texas. The smaller study, by Katsiyannis and Archwamety (1999), examined reincarceration rates within 3 years after release among 549 youth incarcerated in a Nebraska rehabilitation and treatment facility. Both studies compared youth who earned GEDs while incarcerated to those who did not, but only Jeffords and McNitt (1993) adjusted for baseline demographic characteristics, earning a Level 3 on the Maryland Scale, while Katsiyannis and Archwamety (1999) earned a Level 2.

In Table 3, we find that the mean precision-weighted odds ratio across the two studies is 0.534, meaning that those who earned GEDs while incarcerated had postrelease recidivism odds 47 % lower than those who did not. Though the estimate is significant at the 5 % level, selection bias is a particular concern for these studies. We conclude that obtaining a GED in a juvenile correctional facility is associated with lower recidivism, but the extent to which the GED causes this difference remains an open question.

Discussion and Conclusion

Considering What Works

Our synthesis of correctional education interventions for incarcerated juveniles reveals great heterogeneity in terms of interventions, methods, and outcomes of interest. Among the 18 eligible studies we identified, we classified the interventions into five categories: remedial academic education, computer-assisted instruction, personalized instruction, vocational education, and GED completion. Studies that focused on academic outcomes were typically more rigorous but smaller than those that focused on recidivism. Diploma completion and employment were dependent variables in only two studies each.

With such programmatic and methodological variation among a small number of studies, it is difficult to extrapolate clear lessons about what works in juvenile correctional education. Having said that, it is useful to consider in which types of programs and for which outcomes we have the most compelling evidence of positive effects. Limiting our discussion just to estimates that are statistically significant and for which there is a strong basis for causal inference (at least one Level 5 study and no Level 2 studies), we can identify some promising intervention categories. In particular, we find positive evidence for the effect of computer-assisted instruction on reading comprehension (0.21 of a standard deviation). This estimate is driven mainly by the Read 180 randomized trial (Loadman et al., 2011), which is by far the largest study in the category, but the estimate from the Tune In to Reading randomized trial is almost identical in magnitude (Calderone et al., 2009).

We also find positive evidence for the effect of personalized instruction on diploma completion and postrelease employment, with odds ratios of 3.42 and 1.45, respectively. However, these estimates are based only on a single randomized trial, the NCCD (2009) evaluation of the Avon Park Youth Academy, which also found no effect on recidivism. The Avon Park model included a particularly intensive and multifaceted approach to personalized learning, which involved not only personalized academic instruction, but also vocational learning opportunities, continuity in mentoring between incarceration and aftercare, and a lowered ratio of youth to probation officers in aftercare.

Though we find some statistically significant estimates for interventions in each of the other categories, none are based on studies with a strong basis for causal inference, so we cannot conclude that the interventions caused the outcomes of interest.

Key Insights for Researchers and Program Directors

One lesson from this synthesis is that the field of juvenile correctional education is ripe for larger randomized trials. Loadman et al.'s (2011) Read 180 trial and NCCD's (2009) Avon Park Youth Academy trial suggest that such studies, though challenging to undertake, are feasible. Several of the smaller randomized trials we include here have noted the difficulties of high student turnover in correctional facilities, and of simply gaining permission to undertake research in these facilities (Calderone et al., 2009; Shippen et al., 2012). Such

research efforts will clearly take time to develop and execute. They will ideally be realized through long-term partnerships between researchers and correctional facilities. Because such partnerships take time to establish, there may also be a federal role in galvanizing them. The U.S. Department of Education Institute of Education Science's grant program for supporting research partnerships between school systems and researchers offers one potential model. Guided by such partnerships, facilities can make increasingly evidence-based decisions that not only improve young offenders' prospects but also reduce the social incidence of crime and delinquency.

Also notable is the lack of rigorous quasi-experiments and natural experiments, such as a 2013 study by Aizer and Doyle (2013) that found negative effects of juvenile incarceration by capitalizing on naturally occurring random assignment to harsh judges. Studies that leverage random or nearly random processes that shape students' access to particular programs can often yield causal estimates using only administrative datasets (Barrow & Rouse, 2005). Such designs, which are a mainstay of policy research by economists, have been slow to penetrate research in juvenile correctional settings. Whether this is due to the sensitivity and inaccessibility of administrative data in juvenile settings is unclear, but this, too, is an area in which research-practitioner partnerships may show promise.

Though the evidence base about what works in juvenile correctional education remains incomplete, the existing research does suggest promising directions for future programmatic investments. Program directors who make decisions based on extant evidence can help strengthen the research base by documenting their interventions and outcomes using the most rigorous methods at their disposal.

Acknowledgements

This research was funded through a cooperative agreement with the U.S. Department of Justice's Bureau of Justice Assistance (BJA). We are particularly grateful for the support and guidance of our project officers at BJA, Gary Dennis and Thurston Bryant. This work also benefitted greatly from substantive comments by Gerald Gaes, Lynn Karoly, John Linton, and Susan Turner, and from two anonymous peer reviewers. In addition, we are grateful for the insights of our juvenile scientific review team, Joseph Gagnon, Paul Hirschfield, David Houchins, Kristine Jolivette, and Margaret Shippen, and for research assistance by Lopamudra Das. Any errors are the responsibility of the authors.

References

**Indicates a reference included in the meta-analysis.*

Aizer, A., & Doyle, J. J. (2013). *Juvenile incarceration, human capital and future crime: Evidence from randomly-assigned judges.* Cambridge, MA: Massachusetts Institute of Technology.

*Allen-DeBoer, R. A., Malmgren, K. W., & Glass, M.-E. (2006). Reading instruction for youth with emotional and behavioral disorders in a juvenile correctional facility. *Behavioral Disorders, 32*, 18-28.

*Archwamety, T., & Katsiyannis, A. (2000). Academic remediation, parole violations, and recidivism rates among delinquent youths. *Remedial and Special Education, 21*, 161-170.

Barrow, L., & Rouse, C. E. (2005). Causality, causality, causality: The view of education inputs and outputs from economics (2005-15). Chicago, IL: Federal Reserve Bank of Chicago. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=869440

Bland, J. M., & Altman, D. G. (2000). The odds ratio. *British Medical Journal, 320*, 1468. doi: 10.1136/bmj.320.7247.1468

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy, 76*, 169-217.

Biggs, M. C., Homan, S. P., Dedrick, R. F., Minick, V., & Rasinski, T. (2008). Using an interactive singing software program: A comparative study of struggling middle school readers. *Reading Psychology, 29,* 195-213. doi: 10.1080/02702710802073438

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13*, 4-16.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Hoboken, NJ: Wiley.

Bozick, R., & Dalton, B. (2013). Balancing career and technical education with academic coursework: The consequences for mathematics achievement in high school. *Educational Evaluation and Policy Analysis, 35*, 123-138.

Caldagno, J. C., & Long, B. T. (2008). *The impact of postsecondary remediation using a regression discontinuity approach: Addressing endogenous sorting and noncompliance*. New York: National Center for Postsecondary Research.

*Calderone, C., Bennett, S. V., Homan, S., Dedrick, R. F., & Chatfield, A. (2009). Reaching the hard to reach: A comparison of two reading interventions with incarcerated youth. *Middle Grades Research Journal, 4*(3), 61-80.

Chambers, J. G., Parrish, T. B., & Harr, J. J. (2004). What are we spending on special education in the United States, 1999-2000? Report 1. Washington, DC: American Institutes for Research Center for Special Education Finance.

Davis, L. M., Bozick, R., Steele, J. L., Saunders, J., & Miles, J. N. V. (2013). *Evaluating the effectiveness of correctional education: A meta-analysis of programs that provide education to incarcerated adults.* Santa Monica, CA: RAND Corporation.

Davis, L. M., Steele, J. L., Bozick, R., Williams, M., Turner, S., Miles, J. N. V., Saunders, & J. Steinberg, P. S. (2014). *How effective is correctional education, and where do we go from here? The results of a comprehensive evaluation.* Santa Monica, CA: RAND Corporation.

*DelliCarpini, M. (2010). Building a better life: Implementing a career and technical education program for incarcerated youth. *The Journal of Correctional Education, 61,* 283-295.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials, 7,* 177-187. doi: 10.1016/0197-2456(86)90046-2

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

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Journal of Medicine, 315*, 629-634.

Federal Bureau of Prisons. (2015). *Inmate age*. Washington, DC: Federal Bureau of Prisons, U.S. Department of Justice. Retrieved from http://www.bop.gov/about/statistics/statistics_inmate_age.jsp

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*, 239-256.

Glaze, L. E., & Parks, E. (2012). *Correctional populations in the United States, 2011*. Washington, DC: Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Education.

Guerino, P., Harrison, P. M., & Sabol, W. J. (2011). *Prisoners in 2010.* Washington, DC: Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice.

Hazel, N. (2008). *Cross-national comparison of youth justice*. London, UK: Youth Justice Board for England and Wales.

*Houchins, D. E., Jolivette, K., Krezmien, M. P., & Baltodano, H. M. (2008). A multi-state study examining the impact of explicit reading instruction with incarcerated students. *Journal of Correctional Education, 59*, 65-85.

*Jeffords, C. R., & McNitt, S. (1993). *The relationship between GED attainment and recidivism: An evaluation summary*. Austin, TX: Texas Youth Commission.

*Kane, B. J., & Alley, G. R. (1980). A peer-tutored, instructional management program in computational mathematics for incarcerated, learning disabled juvenile delinquents. *Journal of Learning Disabilities, 13*, 39–42. doi: 10.1177/00222194800 1300308

*Katsiyannis, A., & Archwamety, T. (1999). Academic remediation/achievement and other factors related to recidivism rates among delinquent youths. *Behavioral Disorders, 24*, 93-101.

Kemple, J. J., & Willner, C. J. (2008). Career academies: Long-term impacts on labor market outcomes, educational attainment, and transitions to adulthood. New York, NY: MDRC.

Kontopantelis, E., & Reeves, D. (2010). Metaan: Random-effects meta-analysis. *The Stata Journal, 10,* 395-407.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case design technical documentation* (Version 1.0 [Pilot]). Washington, DC: What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education.

Krezmien, M. P., & Mulcahy, C. A. (2008). Literacy and delinquency: Current status of reading interventions with detained and incarcerated youth. *Reading and Writing Quarterly, 24*, 219-238. doi: 10.1080/10573560701808601

Leone, P. E. (1994). Education services for youth with disabilities in a state-operated juvenile correctional system: Case study and analysis. *The Journal of Special Education, 28,* 43-58.

Lhamon, C. E. (2014). Dear colleague letter: Resource comparability. Washington, DC: Office for Civil Rights, U.S. Department of Education.

Lhamon, C. E., & Gupta, V. (2014). Dear colleague letter: Civil rights in juvenile justice residential facilities. Washington, DC: U.S. Department of Education and U.S. Department of Justice.

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

*Loadman, W. E., Moore, R. J., Ren, W., Zhu, J., Zhao, J., & Lomax, R. G. (2011). *Striving Readers Year 5 project evaluation report: Ohio: An addendum to the Year 4 report*. Columbus, OH: The Ohio State University.

Martorell, F., & McFarlin, I. J. (2008). *Help or hindrance? The effects of college remediation on academic and labor market outcomes*. Santa Monica, CA: RAND Corporation

*Mayer, E. E., & Hoffman, R. A. (1982). A comparison of the effects of individualized instruction with group instruction on the academic achievement and self-concept of youthful offenders. *Journal of Correctional Education, 13*, 11–13.

McGraw Hill Education. (2013). *SRA Corrective Reading*. Columbus, OH: McGraw Hill Education. Retrieved from http://www.mcgraw-hill.co.uk/sra/correctivereading.htm

Mears, D. P., & Aron, L. Y. (2003). Addressing the needs of youth with disabilities in the juvenile justice system: The current state of knowledge. Washington, DC: Urban Institute Justice Policy Center.

*National Council on Crime and Delinquency (NCCD). (2009). *In search of evidence-based practice in juvenile corrections: An evaluation of Florida's Avon Park Youth Academy and STREET Smart Program.* Madison, WI: National Council on Crime and Delinquency.

Neild, R. C., Boccanfuso, C., & Byrnes, V. (2013). The academic impacts of career and technical schools: A case study of a large urban school district. Baltimore, MD: Center for the Social Organization of Schools, Johns Hopkins University.

Office of Juvenile Justice and Delinquency Prevention. (2013, October 30). *Statistical briefing book.* Washington, DC: Office of Juvenile Justice and Delinquency Prevention, U.S. Department of Justice. Retrieved from http://www.ojjdp.gov/ojstatbb/corrections/

Pasternak, R., Portillos, R., & Hoff, H. (1988). Providing an appropriate education to adjudicated and incarcerated juvenile delinquents: The challenge to correctional education administrators. *Journal of Correctional Education, 39*, 154-159.

Project READ. (1978). *To make a difference*. Silver Spring, MD: National Criminal Justice Reference Service.

Ritchey, K. D., & Goeke, J. L. (2006). Orton-Gillingham and Orton-Gillingham-based reading instruction: A review of the literature. *The Journal of Special Education, 40*, 171-183.

*Roos, L. E. (2006). *The effect Of career development on employment and recidivism among juvenile offenders*. Minneapolis, MN: Walden University.

Sander, J. B., Patall, E. A., Amoscato, L. A., Fisher, A., & Funk, C. (2012). A meta-analysis of the effect of juvneile delinquency interventions on academic outcomes. *Children and Youth Services Review, 34*, 1695-1708. doi: 10.1016/j.childyouth.2012.04.005

*Scarlato, M. C., & Asahara, E. (2004). Effects of "corrective reading" in a residential treatment facility for adjudicated youth. *Journal of Direct Instruction, 4*, 211-217.

Scholastic. (n.d.). Scheduling options for *Scholastic Read 180* implementations. New York, NY: Scholastic, Inc. Retrieved from http://teacher.scholastic.com/products/statehomepages/uploadfile/resource/fl_read180_schedulingoptions.pdf

Scientific Learning Corporation. (2004). Improved reading skills by students in the Virginia Department of Correctional Education who used Fast ForWord products. *MAPS for Learning: Educator Reports, 8*(28), 1-5.

Scott-Clayton, J., & Rodriguez, O. (2012). *Development, discouragement, or diversion? New evidence on the effects of college remediation* (18328). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w18328.

Sedlak, A. J., & McPherson, K. S. (2010). Youth's needs and services: Findings from the Survey of Youth in Residential Placement. *Juvenile Justice Bulletin*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention, Office of Justice Programs, U.S. Department of Justice.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Sherman, L. W., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1997). Preventing crime: What works, what doesn't, what's promising: A report to the United States Congress. Washington, DC: U.S. Office of Justice Programs.

*Shippen, M. E., Morton, R. C., Flynt, S. W., Houchins, D. E., & Smitherman, T. (2012). Efficacy of a computer-based program on acquisition of reading skills of incarcerated youth. *Remedial and Special Education, 33*, 14-22. doi: 10.1177/0741932510362512

Sickmund, M., Sladky, A. J., Kang, W., & Puzzanchera, C. (2013). *Easy access to the Census of Juveniles in Residential Placement: 1997-2011*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention, U.S. Department of Justice. Retrieved from http://www.ojjdp.gov/ojstatbb/ezacjrp/asp/selection.asp

*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*, 54–66. doi: 10.1007/BF02654938

*Skonovd, N., Krause, W., & Troy, L. A. (1991). The regional youth educational facility: A promising short-term intensive institutional and aftercare

program for juvenile court wards. In T. L. Armstrong (Ed.), *Intensive interventions with high-risk youths: Promising approaches in juvenile probation & parole* (pp. 395-422). Monsey, NY: Criminal Justice Press.

Slavin, R. E., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best-evidence synthesis. *Reading Research Quarterly, 43*, 290-322. doi: 10.1598/RRQ.43.3.4

Snyder, H. N., & Sickmund, M. (2006). *Juvenile offenders and victims: 2006 national report*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention, Office of Justice Programs, U.S. Department of Justice.

Soe, K., Koki, S., & Chang, J. M. (2000). *Effect of computer-assisted instruction (CAI) on reading achievement: A meta-analysis.* Honolulu, HI: Pacific Resources for Education and Learning. Retrieved from http://files.eric.ed.gov/fulltext/ED443079.pdf

Torgesen, J., Myers, D., Schirm, A., Stuart, E., Vartivarian, S., Mansfield, W., … & Haan, C.. (2006). *National assessment of Title I interim report--Volume II: Closing the reading gap: First year findings from a randomized trial of four reading interventions for striving readers*. Washington, DC: Institute

of Education Sciences, U.S. Department of Education. Retrieved from http://www2.ed.gov/rschstat/eval/disadv/title1interimreport/vol2.pdf

Tyler, J. H., & Kling, J. R. (2007). Prison-based education and re-entry into the mainstream labor market. In S. Bushway, M. Stoll, & D. Weiman (Eds.), *Barriers to reentry? The labor market for released prisoners in post-industrial America* (pp. 227-256). New York, NY: Russell Sage Foundation Press.

Tyler, J. H., Murnane, R. J., & Willett, J. B. (2000). Estimating the labor market signaling value of the GED. *The Quarterly Journal of Economics, 115*, 431-468.

U.S. Census Bureau. (2011). *Current population survey, annual social and economic supplement* (Table 1. Population: 2011). Washington, DC: U.S. Census Bureau. Retrieved from https://www.census.gov/population/age/data/2011comp.html

U.S. Census Bureau. (2013, June 27). *State and county quick facts: USA*. Washington, DC: U.S. Census Bureau. Retrieved from http://quickfacts.census.gov/qfd/states/00000.html

U.S. Department of Education. (2013). *National assessment of career and technical education: Interim report*. Washington, DC: Policy and Program

Studies Service, Office of Planning, Evaluation, and Policy Development, U.S. Department of Education.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46,* 197-221. doi: 10.1080/00461520.2011.611369

Wexler, J., Pyle, N., Flower, A., Williams, J. L., & Cole, H. (2014). A synthesis of academic interventions for incarcerated adolescents. *Review of Educational Research, 84,* 3-46. doi: 10.3102/0034654313499410

What Works Clearinghouse. (2007). *Corrective Reading (WWC intervention report)*. Washington, DC: Institute of Education Sciences, U.S. Department of Education.

What Works Clearinghouse. (2009). *Read 180 (WWC intervention report)*. Washington, DC: Institute of Education Sciences, U.S. Department of Education.

What Works Clearinghouse. (2010). *Orton-Gillingham-based strategies (unbranded) (WWC intervention report)*. Washington, DC: Institute of Education Sciences, U.S. Department of Education.

What Works Clearinghouse. (2013). *Fast ForWord (WWC intervention report)*. Washington, DC: Institute of Education Sciences, U.S. Department of Education.

What Works Clearinghouse. (2014). *Procedures and standards handbook (Version 3.0)*. Washington, DC: Institute of Education Sciences, U.S. Department of Education.

Wilson, D. (2011). *Calculating effect sizes*. Oslo, Norway: Campbell Collaboration. Retrieved from http://www.campbellcollaboration.org/artman2/uploads/1/2_D_Wilson__Calculating_ES.pdf

*Wilson, P. R. (1994). Recidivism and vocational education. *Journal of Correctional Education, 45*, 158-163.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Table 1. Operational Definitions of Evidence Ratings on the What Works Clearinghouse and Maryland Scientific Methods Scales

What Works Clearinghouse (WWC) Scale	Maryland (MD) Scientific Methods Scale	Joint Operational Definition
Meets Standards	5	Randomized, controlled trial with attrition below the liberal WWC threshold, *or* single-case designs with well-established pre- and post-intervention trends.
Meets Standards with Reservations	4	Quasi-experimental design (or high-attrition RCT) in which the treatment and comparison groups are matched (within about 1/20th of a standard deviation) at baseline on at least age, prior offenses, baseline educational level, and time to data collection. *Or* single-case designs with moderately establish trends.
Does Not Meet Standards	3	Treatment and comparison groups are matched on 1-2 variables other than gender, and/or there are statistical controls for at least some baseline differences between groups other than gender.
	2	No random assignment or matching, and no statistical controls for baseline differences between treatment and comparison groups.
	1	No separate comparison group.

Table 2. Overall Meta-Analytic Estimates for Each Outcome Type, By Maryland Scale Rating

Category	# of Effects	# of Studies	Wtd. Mean	95% Conf. Interval		Cochrane's Q (heterogeneity test)	p on heterog. test	Egger pub. bias coeff.	Pub. bias p-value
Reading or Math (SD)									
MD 2, 3, or 5	14	10	0.206*	0.129	0.283	5.96	0.744	0.141	0.698
MD 3 or 5	14	10	0.206*	0.129	0.283	5.96	0.744	0.141	0.698
MD 5	9	6	0.206*	0.125	0.287	3.16	0.675	0.063	0.897
Diploma Completion (OR)									
MD 2, 3, or 5	2	2	3.105*	2.385	4.041	0.97	0.325	-	0.317
MD 3 or 5	1	1	3.420*	2.468	4.740				
MD 5	1	1	3.420*	2.468	4.740				
Employment (OR)									
MD 2, 3, or 5	2	2	1.403*	1.161	1.694	0.07	0.798	-	0.317
MD 3 or 5	2	2	1.403*	1.161	1.694	0.07	0.798	-	0.317
MD 5	1	1	1.450*	1.056	1.992				
Recidivism (OR)									
MD 2, 3, or 5	8	8	0.705	0.495	1.004	39.73	0.000	-1.885	0.387
MD 3 or 5	3	3	0.950	0.738	1.223	3.02	0.221	-2.107	0.273
MD 5	1	1	1.042	0.775	1.401				

* $p<.05$ for weighted mean estimates

Notes: SD=standard deviation units; OR=odds ratios. The null hypothesis on the heterogeneity test is that effects are homogeneous across studies. The null hypothesis on the publication bias test is that there is no publication bias. Publication bias p-values for diploma completion and employment are based on Begg rather than Egger tests due to a very small number of estimates.

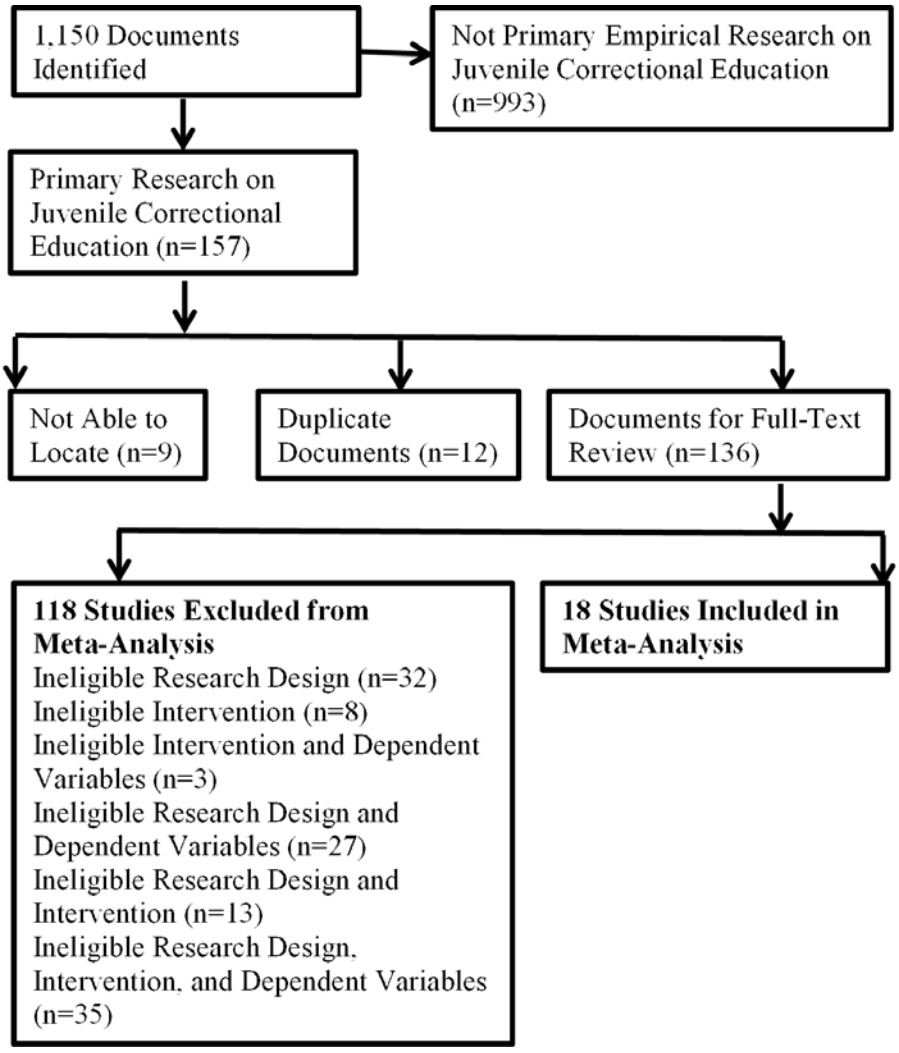
Table 3. Meta-Analytic Estimates for Each Intervention Category, by Outcome Type

Intervention Type	Dependent Variable	# of Effects	# of Studies	Wtd. Mean	95% Conf. Interval		Q	p on heterog. test	Basis for Causal Inference
Remedial	*Word-level reading (SD)*	3	3	0.534	-0.388	1.456	0.34	0.843	*High:* 2 level 5, 1 level 3
	Reading comprehension (SD)	2	2	0.526*	0.056	0.996	0.05	0.823	*Moderate:* 2 level 3
	Recidivism (OR)	2	2	0.936	0.210	4.162	6.98	0.008	*Lower:* 1 level 3, 1 level 2
CAI	*Word-level reading (SD)*	1	1	-0.123	-0.673	0.427			*High:* 1 level 5
	Reading comprehension (SD)	3	3	0.201*	0.120	0.283	1.79	0.409	*High:* 3 level 5
Personalized	*Reading/Math (SD)*	5	3	0.145	-0.129	0.420	1.64	0.439	*High:* 3 level 5, 2 level 3
	Diploma Complet. (OR)	1	1	3.420*	2.468	4.740			*High:* 1 level 5
	Employment (OR)	1	1	1.450*	1.056	1.992			*High:* 1 level 5
	Recidivism (OR)	2	2	0.577	0.138	2.424	4.3	0.038	*Moderate:* 1 level 5, 1 level 2
Vocational	*Diploma Complet. (OR)*	1	1	2.59*	1.657	4.050			*Lowest:* 1 level 2
	Employment (OR)	1	1	1.377*	1.089	1.742			*Moderate:* 1 level 3
	Recidivism (OR)	2	2	0.669	0.307	1.459	8.86	0.003	*Lower:* 1 level 3, 1 level 2
GED Completion	*Recidivism (OR)*	2	2	0.534*	0.422	0.677	0.97	0.326	*Lower:* 1 level 3, 1 level 2

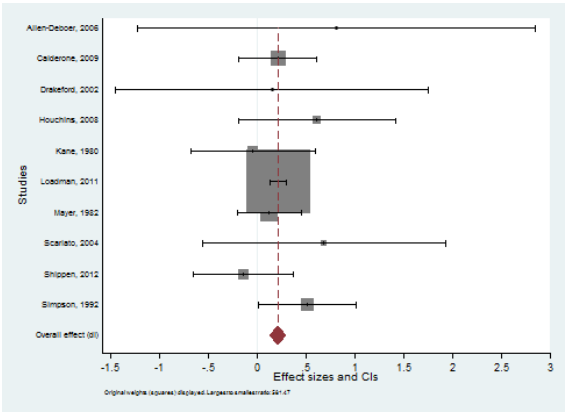
* p<.05 for weighted mean estimates

Notes: SD=standard deviation units; OR=odds ratios. The null hypothesis on the heterogeneity test is that effects are homogeneous across studies.

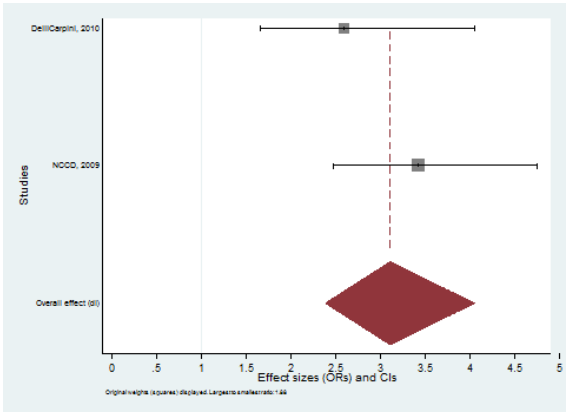
Figure 1. Screening Process for Manuscript Eligibility



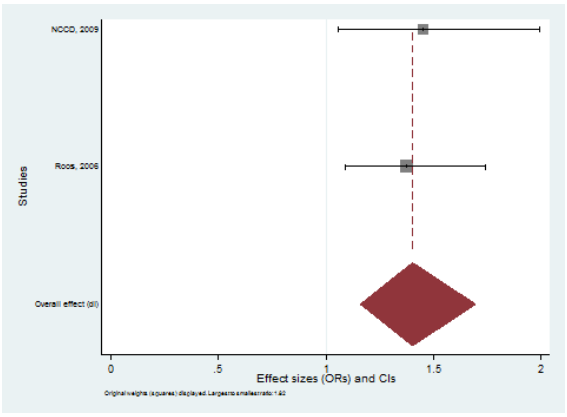
Appendix Figure A1. Forest Plots with Overall Meta-Analytic Estimates and Confidence Intervals for the Four Dependent Variables of Interest



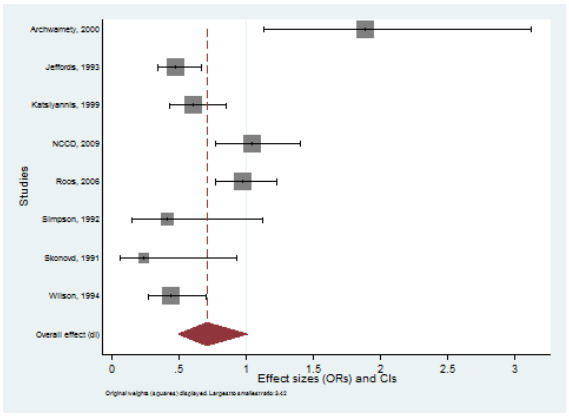
Academic Performance (SD Units)



Diploma Completion (Odds Ratios)

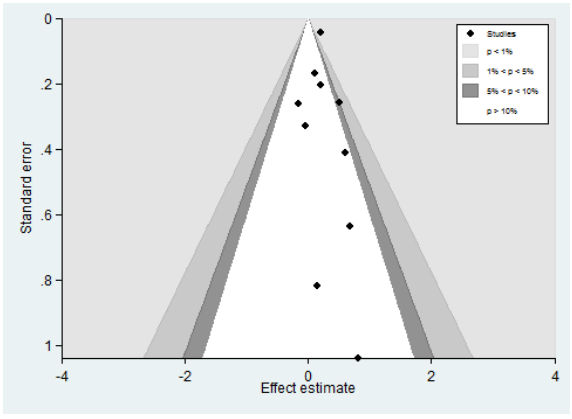


Employment (Odds Ratios)

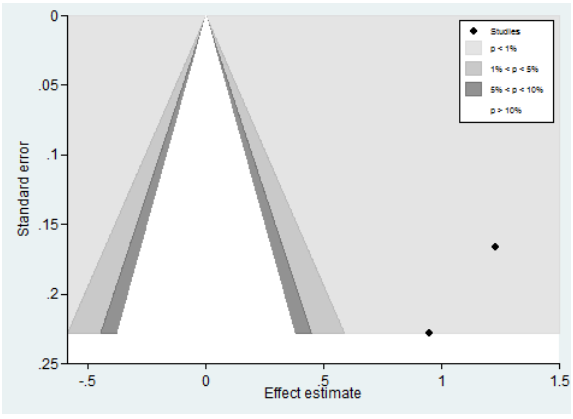


Recidivism (Odds Ratios)

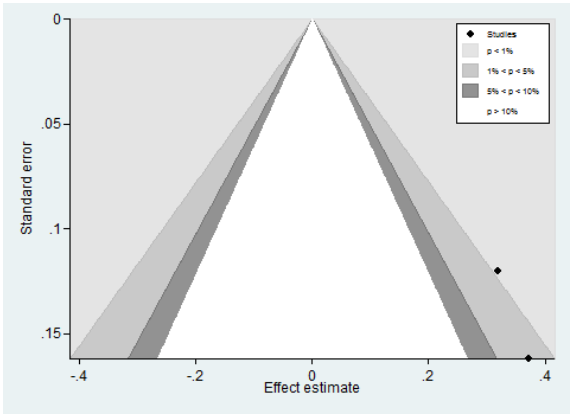
Appendix Figure A2. Funnel Plots of Overall Meta-Analytic Estimates for the Four Dependent Variables of Interest



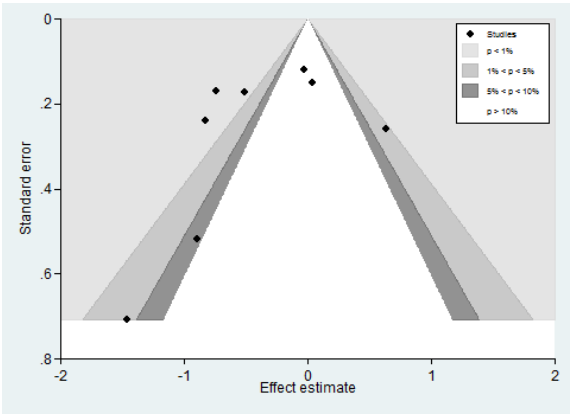
Academic Performance (SD Units)



Diploma Completion (Log Odds Ratios)



Employment (Log Odds Ratios)



Recidivism (Log Odds Ratios)

Appendix Table A1. Details about Studies, Samples, and Effect Estimates in the Meta-Analysis

Citation	Treatment Condition	Comparison Condition	Setting	Demographics	N_T	N_C	Duration and Frequency	Test Score Effects	Diploma Completion/ Employment Effects	Recidivism Effects	MD Scale
Remedial Academic Instruction											
Allen-DeBoer et al., 2006	Corrective Reading	Traditional language arts instruction	Mental health treatment unit within a juvenile correctional facility	Age: 16–18; 100% Male; 75% African American; 25% White; 100% with learning disabilities; Baseline skills: 4th-5th grade	4	0	30 minutes a day, 5 days a week, for 9 weeks (30 lessons on average)	Words Read Correctly per Minute, mean change: 35.75 (0.806 SD, no hypothesis test)			5
Drakeford, 2002	Corrective Reading	Traditional language arts instruction	Oak Hill Academy in Maryland	Age: 12–21 (mean: 17); 100% Male; 100% African American; 100% with history of educational disabilities	6	0	One hour, 3 times a week, for 8 weeks (20 lessons on average)	Words Read Correctly per Minute, mean change: 9.17 WPM (0.148 SD, no hypothesis test)			5
Scarlato and Asahara, 2004	Corrective Reading for (180 minutes per week)	Reading Specialist for 60 minutes twice a week, plus 225 minutes of additional reading instruction (345 min/wk)	Residential juvenile treatment facility	Age: 16–17; 100% Male; 100% with learning disabilities or emotional disturbance; 100% read below grade level	5	4	45 minutes, 4 times a week for 19 weeks	Woodcock Reading Mastery,: Basic Skills: 0.69 SD (p>.05); Reading Comprehension: 0.67 SD (p>.05)			3
Simpson, Swanson, and Kunkel, 1992	Orton-Gillingham structured remedial reading instruction for 90 minutes a day in groups of 1–6	Default language arts instruction for 45 minutes a day in classes of about 12	Two juvenile youth detention facilities (location not given)	Age: 13–18; 100% Male; Baseline reading grade level: 4.4; 100% learning disabled	32	31	Actual mean dosage: 51.9 hours in treatment group vs. 46.0 in control group	Woodcock Reading Mastery: 0.86 years, or 0.51 SD (p=.007)		Re-arrest within a year following release: -22 percentage points (p=.015) T: 41% C: 63%	3

Citation	Treatment Condition	Comparison Condition	Setting	Demographics	N_T	N_C	Duration and Frequency	Test Score Effects	Diploma Completion/ Employment Effects	Recidivism Effects	MD Scale
Archwamety and Katsiyannis, 2000	Remedial education in math or reading	Non-remedial education	Nebraska Youth Rehabilitation and Treatment Center	Age: 12–18; Mean IQ: 94.3; treatment students were at least one grade behind in remedial subject	339	166	Not specified			Recidivism within 1–7 years after release: +9.4 pctg. points (p<.05) T: 23.3% C: 13.9%	2
Computer-Assisted Instruction											
Loadman et al., 2011	Read 180 (Scholastic)	Default English language arts instruction	Eight Ohio Department of Youth Services facilities	Age: 14–22, most in grades 9–10; 96% Male; 69% African American; 24% White; 2% Hispanic; 5% Other; 48% with disabilities; 100% baseline reading level at least basic but below proficient	677	568	90 min., 5 days a week, for 20 weeks	Scholastic Reading Inventory (SRI) score: 0.21 SD (p<.001)			5
Shippen et al, 2012	Fast ForWord software-based beginning reading program (Scientific Learning)	Default, individualized academic and vocational training	Long-term maximum security juvenile facility in Alabama	Age: 11–20 (mean=16.3); 100% Male 53% African American; 45% White; 2% Other; Mean IQ: 78; 18% with mild learning disabilities	27	24	45 min., 5 days a week, for 11 weeks (average=24 days)	Test of Word Reading Efficiency: -0.123 SD (p>.05) Woodcock Reading Mastery: -0.17 SD (p>.05)			5

Citation	Treatment Condition	Comparison Condition	Setting	Demographics	N_T	N_C	Duration and Frequency	Test Score Effects	Diploma Completion/ Employment Effects	Recidivism Effects	MD Scale
Calderone et al, 2009	Tune in to Reading (TIR), a program to teach reading through singing (Electronic Learning Products)	Default instructional program (namely, FCAT Explorer, an online, standards-based program)	Six residential sites for juveniles in the Florida correctional system	Ages not given; grades 7–11; 100% Male 52% African American 13% Hispanic 31% White 44% with disabilities	64	39	45 minutes, twice a week, for 9 weeks	Computer-adaptive cloze reading assessment, 0.206 SD (p>.05)			5
Personalized Academic Instruction											
National Council on Crime and Delinquency, 2009	Avon Park Youth Academy: Intensive, personalized, vocational and academic training with aftercare	Default juvenile correctional programs within the state	Florida Department of Juvenile Justice facilities	Age: 16–18; 41% African American; 14% Hispanic; 44% White; 38% with special needs; 65% reading below grade 6; 100% with math skills below grade 6	369	345	14.2 month average stay in facility (versus 11.2 months for comparison group)		High school or GED completion by release: 27.1 pctg. points (p<.01) T: 49.1% C: 22.0% Employment 1 year post-release: 8 pctg. points (p=.02) T: 72.4% C: 64.4%,	Re-arrest within a year: 1.0 pctg. point (p>.2) T: 57.2% C: 56.2%	5
Skonovd et al., 1991	Intensive, competency-based education with vocational training and aftercare	Default programs for juveniles in the same county	San Bernardino County Probation Department Juvenile Hall	Age: 16–17; 21% African American; 29% Hispanic; 50% White	25	20	6 months in juvenile facility and 4–6 months in after care			Re-arrest or probation violation within 6 months: -29 pctg. points (p<.05) T: 16%; C: 45%	2
Mayer and Hoffman, 1982	Individualized academic instruction	Group (classroom-level) instruction	Four youth offender facilities in Florida	Ages not given; 100% Male 52% African American 48% White	68	75	10 months (frequency not given)	California Achievement Test, (math, reading, language): 2 months of learning (0.118 SD, no hypoth. test)			3

Citation	Treatment Condition	Comparison Condition	Setting	Demographics	N_T	N_C	Duration and Frequency	Test Score Effects	Diploma Completion/ Employment Effects	Recidivism Effects	MD Scale
Kane and Alley, 1980	Peer-managed instruction with tutor-student ratio of 1:1 to 1:2	Teacher-managed instruction with teacher- student ratio of 1:3 to 1:7	Minimum-security juvenile institution in Minnesota	Age: 12–17; 100% learning disabled; mean pretest math grade level: 6.0	21	17	8 weeks (38 45- minute class periods)	Science Research Associates Mathematics Assessment: -0.045 SD (p>.05)			3
Houchins et al, 2008	Corrective Reading: 1:4 teacher-student ratio	Corrective Reading: 1:12 teacher-student ratio	Long-term juvenile correction facility in a Mid-Atlantic State	Age: 13–17 (mean: 16.5); 100% Male; 64% African American; 18% Hispanic; 18% White; 21% with learning/cognitive disabilities; 58% with emotional or behavioral disabilities	10	10	1 hour, 3 times a week, for 7 weeks (21 sessions)	Woodcock Reading Mastery, Word Identification: 0.60 SD (p<.01); Word Attack: 0.50 SD; (p<.01) Gray Silent Reading: 0.72 SD (p<.01)			5
Vocational Education/CTE											
Roos, 2006	Re-Integration of Offenders–Youth (RIO-Y) career development course	No participation in a career development course	Texas Youth Commission facilities	Age: 18–21; 34% African American; 38% Hispanic; 28% White	582	920	30 days of instruction (versus none in comparison group)		Employment 1 year post-release, odds ratio: 1.39 (p<.01)	Re-arrest within 1 year, odds ratio: 0.97 (p=.8)	3
Wilson, 1994	Vocational education elective participation in facility (auto, business, construction, food, special services)	Participation in nonvocational education	Colorado Division of Youth Services facilities	Age: 11–18; 100% Male; 16% Black; 34% Hispanic; 48% White; 2% Other	159	143	Not reported			Reincarceration within 5 years: -17.1 pctg. points (p<.05) T: 61.2% C: 78.3%	2

Citation	Treatment Condition	Comparison Condition	Setting	Demographics	N_T	N_C	Duration and Frequency	Test Score Effects	Diploma Completion/ Employment Effects	Recidivism Effects	MD Scale
DelliCarpini, 2010	Vocational education program availability (business, drafting, carpentry)	Participation in default educational program	Eastern Suffolk BOCES Program for Incarcerated Youth in NY State	Age: 16–21; No additional demographic information provided	465	581	8 week module (daily instruction implied)		GED pass rate: 7.6 percentage points (p<.001) T: 13.1% C: 5.5%		2
GED Completion											
Jeffords and McNitt, 1993	GED completion in facility	No GED completion in facility	Texas Youth Commission or Gulf Coast Trades Center programs	Age: 16–21; No additional demographic information provided	475	1242	Not reported			Reincarceration within 1 year: -5.8 pctg. points (p<.01)	3
Katsiyannis and Archwamety, 1999	GED completion in facility	No GED completion in facility	Youth rehabilitation and treatment facility in Nebraska	Age: 12–18; 100% Male	284	265	At least 4 months spent in facility			Reincarceration within 3 years: -12.5 pctg. points (p<.01) T: 47.5% C: 60.0%	2