This report is part of the RAND Corporation research report series. RAND reports present research findings and objective analysis that address the challenges facing the public and private sectors. All RAND reports undergo rigorous peer review to ensure high standards for research quality and objectivity.

Evaluating the Effectiveness of Correctional Education

A Meta-Analysis of Programs That Provide Education to Incarcerated Adults

Lois M. Davis, Robert Bozick, Jennifer L. Steele, Jessica Saunders, Jeremy N. V. Miles

Sponsored by the Bureau of Justice Assistance

RAND CORPORATION

BJA
Bureau of Justice Assistance
U.S. Department of Justice

Evaluating the Effectiveness of Correctional Education

A Meta-Analysis of Programs That Provide Education to Incarcerated Adults

Lois M. Davis, Robert Bozick, Jennifer L. Steele, Jessica Saunders, Jeremy N. V. Miles

Sponsored by the Bureau of Justice Assistance

The research described in this report was sponsored by the Bureau of Justice Assistance and conducted in the Safety and Justice Program within RAND Justice, Infrastructure, and Environment.

This project was supported by Grant No. 2010-RQ-BX-001 awarded by the Bureau of Justice Assistance to the RAND Corporation. The Bureau of Justice Assistance is a component of the Office of Justice Programs, which also includes the Bureau of Justice Statistics, the National Institute of Justice, the Office of Juvenile Justice and Delinquency Prevention, the Office for Victims of Crime, and the Office of Sex Offender Sentencing, Monitoring, Apprehending, Registering, and Tracking. Points of view or opinions in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

Library of Congress Cataloging-in-Publication Data is available for this publication.

ISBN: 978-0-8330-8108-7

Cover photo courtesy of PrisonEducation.com.

RAND OFFICES

SANTA MONICA, CA • WASHINGTON, DC

PITTSBURGH, PA • NEW ORLEANS, LA • JACKSON, MS • BOSTON, MA

DOHA, QA • CAMBRIDGE, UK • BRUSSELS, BE

www.rand.org

Foreword

Each year, thousands of incarcerated adults leave the nation's prisons and jails and return to their families and communities. While many successfully reintegrate into their communities, find jobs, and become productive members of society, many others will commit new crimes and end up being reincarcerated. Although a number of factors account for why some ex-prisoners succeed and some don't, we know that a lack of education and skills is one key reason. This is why correctional education programs—whether academically or vocationally focused—are a key service provided in correctional facilities across the nation. But do such correctional education programs actually work? We care about the answer both because we want ex-prisoners to successfully reenter communities and because we have a responsibility to use taxpayer dollars judiciously to support programs that are backed by evidence of their effectiveness—especially during difficult budgetary times like these. Across this Administration, we are committed to investing in evidence-based programming, investigating promising practices, and making science a priority.

Fortunately, the passage of the Second Chance Act of 2007 gave us a chance to comprehensively examine the effectiveness of correctional education because it includes a specific provision to improve education in U.S. prisons and jails. The Bureau of Justice Assistance, with guidance from the Office of Vocational and Adult Education, competitively awarded a project to the RAND Corporation in 2010. We asked RAND to comprehensively examine the current state of correctional education for incarcerated adults and juveniles and where the field is headed, which correctional education programs are effective, and how effective programs can be implemented across different settings. This valuable report—a new meta-analysis examining the effectiveness of correctional education programs—is a key part of that effort and can help us answer the question of whether the nation's investment in correctional education is indeed achieving its intended outcomes.

The results presented here are truly encouraging. Confirming the results of previous meta-analyses—while using more (and more recent) studies and an even more rigorous approach to selecting and evaluating them than in the past—RAND researchers show that correctional education reduces postrelease recidivism and does so cost-effectively. And the study also looks at another outcome key to successful reentry—postrelease employment—and finds that correctional education may increase such employment. The reason the findings for employment are merely suggestive is that only one of the 19 studies that evaluated post-employment outcomes used a highly rigorous methodology.

This need for more high-quality studies that would reinforce the findings is one of the key areas the study recommends for continuing attention. Just as important is the need to better

understand what makes some programs more effective than others—is it the program design, the type of instruction, the length of the program, or, more likely, some combination of these and other factors? Having such knowledge is key to telling us which programs should be developed and funded—which programs will provide the greatest return on taxpayer dollars. Other parts of the RAND project, including an assessment of best practices derived from examining current programs, will further illuminate what works, but new and ongoing studies should be designed in ways that help isolate the causal effects of particular program designs.

The results provided here give us confidence that correctional education programs are a sound investment in helping released prisoners get back on their feet—and stay on their feet—when they return to communities nationwide. We are pleased to have been able to work cooperatively across our two agencies with the RAND staff and to offer this important information.

Denise E. O'Donnell, J.D.
Director, Bureau of Justice Assistance
Office of Justice Programs
U.S. Department of Justice

Brenda Dann-Messier, Ed.D.
Assistant Secretary
Vocational and Adult Education
U.S. Department of Education

Preface

The Second Chance Act of 2007 (Public Law 110-199) represented a historic piece of legislation designed to improve outcomes for and provide a comprehensive response to the increasing number of individuals who are released from prisons, jails, and juvenile residential facilities and returning to communities upon release. The Second Chance Act's grant programs are funded and administered by the Office of Justice Programs within the U.S. Department of Justice. In 2010, for the first time, funding was set aside for a comprehensive study of correctional education. The Office of Justice Programs' Bureau of Justice Assistance awarded the RAND Corporation a cooperative agreement to undertake a comprehensive examination of the current state of correctional education for incarcerated adults and juveniles and where it is headed, which correctional education programs are effective, and how effective programs can be implemented across different settings. One key task was to undertake a comprehensive review of the scientific literature and a meta-analysis to synthesize the findings from multiple studies as to the effectiveness of correctional education programs in helping to reduce recidivism and improve postrelease employment outcomes. In this report, we detail the meta-analytic approach and findings for academic programs and vocational training programs provided to incarcerated adults. In a subsequent report, we will present the findings for the overall project.

These results will be of interest to federal and state policymakers; administrators of state departments of corrections, public safety, and education; correctional as well as community college educators; career technical training providers; and other organizations that provide educational services and training to currently incarcerated or formerly incarcerated adults. These results will also be of interest to those in the U.S. Departments of Justice and Education who are committed to ensuring the availability and quality of correctional education programs for incarcerated adults.

The RAND Safety and Justice Program

The research reported here was conducted in the RAND Safety and Justice Program, which addresses all aspects of public safety and the criminal justice system, including violence, policing, corrections, courts and criminal law, substance abuse, occupational safety, and public integrity. Program research is supported by government agencies, foundations, and the private sector.

This program is part of RAND Justice, Infrastructure, and Environment, a division of the RAND Corporation dedicated to improving policy and decisionmaking in a wide range of

policy domains, including civil and criminal justice, infrastructure protection and homeland security, transportation and energy policy, and environmental and natural resource policy.

Questions or comments about this report should be sent to the project leaders, Lois M. Davis, Ph.D. (Lois_Davis@rand.org) and Robert Bozick, Ph.D. (Robert_Bozick@rand.org). For more information about the Safety and Justice Program, see http://www.rand.org/safety-justice or contact the director at sj@rand.org.

Contents

Figures

Tables

Summary

Introduction

It is challenging to prepare offenders with the needed vocational skills and education to be successful in reintegrating back into society. Offenders, on average, are less educated than the general population. For example, in 2004, approximately 36 percent of individuals in state prisons had attained less than a high school education compared with 19 percent of the general U.S. population age 16 and over. In addition to having lower levels of educational attainment, offenders often lack vocational skills and a steady history of employment, which is a significant challenge for individuals returning from prison to local communities. And the dynamics of prison entry and reentry make it hard for this population to accumulate meaningful, sustained employment experience. Finally, the stigma of having a felony conviction on one's record is a key barrier to postrelease employment.

On April 9, 2008, the Second Chance Act (Public Law 110-199) (SCA) was signed into law. This important piece of legislation was designed to improve outcomes for individuals who are incarcerated, most of whom will ultimately return to communities upon release. The SCA's grant programs are funded and administered by the Office of Justice Programs (OJP) within the U.S. Department of Justice (DOJ). In 2010, funding was set aside, for the first time under the SCA, to conduct a comprehensive study of correctional education. OJP's Bureau of Justice Assistance (BJA) awarded the RAND Corporation a cooperative agreement to comprehensively examine the current state of correctional education for incarcerated adults and juveniles and where it is headed, which correctional education programs are effective, and how effective programs can be implemented across different settings. One central task in that effort was to comprehensively review the scientific literature and conduct a meta-analysis to synthesize the findings from multiple studies about the effectiveness of correctional education programs in helping to reduce recidivism and improve employment outcomes for incarcerated adults within U.S. state prisons.

In this report, we present the findings from our meta-analysis, which will inform policymakers, educators, and correctional education administrators interested in understanding the association between correctional education and reductions in recidivism and improvements in employment and other outcomes.

To prepare for the meta-analysis, we first conducted a comprehensive literature search for published and unpublished studies released between 1980 and 2011 that examined the relationship between correctional education participation and inmate outcomes. We focused exclusively on studies published in English of correctional education programs in the United States that included an academic and/or vocational curriculum with a structured instructional component. A scientific review panel abstracted data, and the quality of the research design

was rated using the Maryland Scientific Methods Scale and the U.S. Department of Education's What Works Clearinghouse rating scheme. Studies that met our eligibility criteria in terms of intervention type, research design, and outcomes and that rated a 2 or higher on the Maryland Scientific Methods Scale were included in the meta-analysis.

We used meta-analytic techniques to synthesize the effects of correctional education programs administered to adults across multiple studies. As with previous meta-analyses in this area, our focus was largely on recidivism, because it is the outcome most often used in the literature. However, we also examined whether participating in a correctional education program was associated with an increase in labor force participation and whether participating in a correctional education program with a computer-assisted instructional component was associated with gains in achievement test scores. In addition, we conducted a cost analysis comparing the direct costs of correctional education with those of re-incarceration to place our recidivism findings into a broader context.

Results

Relationship Between Correctional Education Programs and Recidivism

Our meta-analytic findings provide additional support for the premise that receiving correctional education while incarcerated reduces an individual's risk of recidivating after release. After examining the higher-quality research studies, we found that, on average, inmates who participated in correctional education programs had *43 percent lower odds of recidivating* than inmates who did not. These results were consistent even when we included the lower-quality studies in the analysis. This translates into a reduction in the risk of recidivating of 13 percentage points for those who participate in correctional education programs versus those who do not. This reduction is somewhat greater than what had been previously reported by Wilson, Gallagher, and MacKenzie (2000), which showed an average reduction in recidivism of about 11 percentage points. Using more recent studies and ones of higher quality, our findings complement the results published by Wilson, Gallagher, and MacKenzie (2000), Aos, Miller, and Drake (2006), and MacKenzie (2006) and provides further support to the assertion that correctional education participants have lower rates of recidivism than nonparticipants.

Given the high percentage of state prison inmates who have not completed high school, participation in high school/general education development (GED) programs was the most common approach to educating inmates in the studies we examined. Focusing only on studies that examined this kind of program relative to no correctional education, we found that inmates who participated in high school/GED programs had *30 percent lower odds* of recidivating than those who had not. In general, studies that included adult basic education (ABE), high school/GED, postsecondary education, and/or vocational training programs showed a reduction in recidivism. However, we could not disentangle the effects of these different types of educational programs, because inmates could have participated in multiple programs, and the amount of time that they spent in any given program was rarely reported.

Relationship Between Correctional Education Programs and Employment

When we look at the relationship between correctional education and postrelease employment, our meta-analyses found—using the full set of studies—that *the odds of obtaining employment postrelease among inmates who participated in correctional education (either academic or vocational*

programs) was 13 percent higher than the odds for those who had not participated. However, only one study fell into the higher-quality category. Thus, if policymakers want to base decisions on the higher-quality studies alone, then we are limited in our ability to detect a statistically significant difference between program participants and nonparticipants in postrelease employment. Still, our results suggest a positive association between correctional education and postrelease employment. Our findings align with those produced in the Wilson, Gallagher, and MacKenzie (2000) meta-analysis, which also found improved odds of employment among correctional education participants.

When examining the relationship between correctional education and postrelease employment, one might expect vocational training programs to be more adept than academic education programs at imparting labor market skills, awarding industry-recognized credentials, or connecting individuals with prospective employers. And, indeed, when we looked at the relationship between vocational training—versus academic correctional education programs—and postrelease employment, we found that *individuals who participated in vocational training programs had odds of obtaining postrelease employment that were 28 percent higher than individuals who had not participated.* In comparison, individuals who participated in academic programs (combining ABE, high school/GED, and postsecondary education programs) had only 8 percent higher odds of obtaining postrelease employment than those individuals who had not participated in academic programs. Although the results suggest that vocational training programs have a greater effect than academic programs on one's odds of obtaining postrelease employment, there was no statistically significant difference between the odds ratios for the two types of programs, because the number of vocational training studies was relatively small.

Relationship Between Computer-Assisted Instruction and Academic Performance

We also examined the relationship between computer-assisted instruction and academic performance. In this case, the outcomes of interest were standardized test scores in mathematics or reading. We reviewed four studies that compared the achievement test scores of inmates receiving computer-assisted instruction with the achievement test scores of inmates receiving face-to-face instruction. In two of the studies, students in both the treatment and comparison groups also received additional, traditional classroom instruction beyond the portion of their instructional time that was computer-assisted. *We estimated that the overall effect of computer-assisted instruction relative to traditional instruction is 0.04 grade levels in reading, or about 0.36 months of learning, and 0.33 grade levels in mathematics, which represents about 3 months of learning.* In other words, on average across the studies, students exposed to computer-assisted instruction relative to traditional instruction learned very slightly more in reading in the same amount of instructional time and substantially more in mathematics. However, there was no statistically significant difference in test scores between the different methods of instruction, and given that the confidence intervals included zero for both reading and mathematics, we could not rule out the possibility that the effects estimated were due to chance alone. Because computer-assisted instruction can be self-paced and supervised by a tutor or an instructor, it is potentially less costly to administer. It is worth noting that, since the publication of these four studies, the capability and utility of instructional technology has progressed substantially (U.S. Department of Education, 2010), which suggests that the effects of the newer technologies may potentially outstrip those found in the studies examined here.

Comparison of the Costs of Correctional Education Programs and Reincarceration Costs

State policymakers, corrections officials, and correctional education administrators are asking a key question: How cost-effective is correctional education? Our cost analysis suggests that correctional education programs are cost-effective. Focusing only on the *direct costs* of correctional education programs and of incarceration itself, and using a three-year reincarceration rate for a hypothetical pool of 100 inmates, we estimated that the three-year reincarceration costs for those who did not receive correctional education would be between $2.94 million and $3.25 million. In comparison, for those who did receive correctional education, the three-year reincarceration costs would be between $2.07 million and $2.28 million. This means that reincarceration costs are $0.87 million to $0.97 million *less* for those who receive correctional education. In comparison, our estimates indicate that the costs of providing education to inmates would range from $140,000 to $174,400 for the pool of 100 inmates. This translates into a per-inmate cost of correctional education ranging from $1,400 to $1,744, suggesting that providing correctional education is cost-effective compared with the cost of reincarceration. It is worth noting that this estimate takes into account only the direct costs to the system, but it does not consider such other costs as the financial and emotional costs to victims of crime or to the criminal justice system as a whole. Hence, it is a conservative estimate of the broader effect that correctional education can potentially yield.

To further help interpret the cost savings, we also calculated the *break-even point*—defined as the risk difference in the reincarceration rate required for the cost of correctional education to be equal to the cost of incarceration. For a correctional education program to be cost-effective, we estimated that a program would need to reduce the three-year reincarceration rate by between 1.9 and 2.6 percentage points to break even. In fact, as noted, our meta-analytic findings show that participation in correctional education programs is associated with a 13 percentage-point reduction in the risk of reincarceration three years after release from prison.

Conclusions and Recommendations

Our meta-analytic findings provide further support that receiving correctional education while incarcerated reduces an individual's risk of recidivating after release from prison. Our findings were stable even when we limited our analyses to those studies with more rigorous research designs. We found a notable effect across all levels of education, from adult basic education and GED programs to postsecondary and vocational education programs. Further, our cost analysis suggests that correctional education programs can be cost-effective. As noted by other researchers interested in estimating the effect of correctional education (e.g., MacKenzie, 2008; Gaes, 2008), we, too, found a number of methodological weaknesses in the current body of research that substantially limit one's ability to inform the direction of policy and the design of effective programs. Thus, a number of questions of interest to educators and policymakers remain that the current literature does not permit us to answer, such as understanding what is inside the "black box" in terms of what program elements, for example, are associated with effective programs.

In addition, much is changing in the field of correctional education. The 2008 recession affected correctional education (and other rehabilitative) programs in a number of states and led to some dramatic changes in the number of programs offered, the sizes of classes, the

modes of delivery, and the number of inmates who participate in these programs. A reduced funding environment will likely be true for many correctional education programs for the near future, and questions about the return on investment of these programs will likely continue to be a topic in state-level budget discussions.

Going forward, there is a need to undertake studies that "drill down" to get inside the black box and identify the characteristics of effective programs in terms of such variables as curriculum, dosage, and quality. To inform policy and funding decisions at the state and federal levels, policymakers need additional information and a better understanding of how these programs work (or do not work). In addition, we need to continue to build the evidence base in this area. We provide recommendations for doing so in four critical areas: (1) applying stronger research designs, (2) measuring program dosage, (3) identifying program characteristics, and (4) examining more proximal indicators of program efficacy.

One option is for state and federal policymakers and foundations to invest in well-designed evaluations of correctional education programs to inform such policy questions. Also, researchers and program evaluators need to strive to implement rigorous research designs to examine questions related to potential bias and program dosage and to measure both proximal and distal outcomes. Funding grants and guidelines can help further the field by requiring the use of more rigorous research designs. Such funding would also enable correctional educators to partner with researchers and evaluators to undertake rigorous and comprehensive evaluations of their programs. Last, a study registry of correctional education evaluations would help in further developing the evidence base in this field to inform policy and programmatic decisionmaking.

Findings from this study can be found on the project's website: http://www.rand.org/jie/projects/correctional-education.html.

Acknowledgments

We are particularly grateful for the guidance and feedback provided throughout this project by our Bureau of Justice Assistance project officers, Gary Dennis, Senior Policy Advisor for Corrections, and Thurston Bryant, Policy Advisor. We are also grateful for the valuable input and feedback provided by Brenda Dann-Messier, Assistant Secretary for Vocational and Adult Education, and John Linton, Director, Office of Correctional Education, Office of Vocational and Adult Education, U.S. Department of Education. We also appreciate the support and insights provided by Stephen Steurer, Executive Director of the Correctional Education Association.

The overall direction of the project was guided in part by a steering committee that included John Dowdell (Director of the Gill Center for Business and Economic Education at Ashland University and Co-Editor of the Journal of Correctional Education), William Sondervan (Professor and Director of Criminal Justice, Investigative Forensics, and Legal Studies at the University of Maryland University College), Stephen Steurer (Executive Director of the Correctional Education Association), and Susan Turner (Professor of Criminology, Law, and Society, at the University of California–Irvine).

In addition, a number of individuals within and outside RAND contributed to various aspects of the project. The Scientific Review Team members helped guide the selection of intervention characteristics to be abstracted and served as independent reviewers in abstracting the study information that were inputs for the meta-analysis. They included Cathryn Chappell (Ashland University), John Dowdell (Ashland University), Joseph Gagnon (University of Florida), Paul Hirschfield (Rutgers University), Michael Holosko (University of Georgia), David Houchins (Georgia State University), Kristine Jolivette (Georgia State University), Larry Nackerud (University of Georgia), Ed Risler (University of Georgia), and Margaret Shippen (Auburn University).

Without the help of the following people, our study would have not been possible. Staff from the RAND library worked tirelessly to locate and procure all documents needed for our study: Tomiko Envela, Brooke Hyatt, and Sachi Yagyu. A team of doctoral students in the Pardee RAND Graduate School helped organize and review all the studies that were considered for inclusion into our meta-analyses: Nono Ayivi-Guedehoussou, Stephanie Chan, Megan Clifford, Lopamudra Das, Russell Lundberg, Shannon Maloney, Christopher McLaren, and Nicole Schmidt. Certain studies required additional review to ensure that the information was coded properly. This was undertaken by Ph.D.-level research staff at RAND: Ramya Chari, Sarah Greathouse, Lisa Sontag-Padilla, Vivian Towe, and Malcolm Williams. Susanne Hempel and Becky Kilburn advised the project on systematic review procedures. Sue Phillips provided website development support for the project, Aneesa Motala assisted with systematic review software support, and Roald Euller provided programming support. We also wish to

acknowledge the work of Paul Steinberg who served as the communications analyst on this report. Dionne Barnes-Proby provided project management and research assistance, and Judy Bearer provided administrative support. We also benefited from the editing and publications production support provided by James Torr, Patricia Bedrosian, and Jocelyn Lofstrom.

Last, we appreciate the insights provided by our technical reviewers, Juan Saavedra, an associate economist at RAND, and David Wilson, Chair of the Criminology, Law, and Society Department at George Mason University.

Abbreviations

ABE	adult basic education
ABLE	Adult Basic Learning Examination, Level II
AIMS	Advanced Instructional Management System
ASE	adult secondary education
AUTOSKILL	AUTOSKILL Component Reading Subskills Program
BJA	Bureau of Justice Assistance
BJS	Bureau of Justice Statistics
CASAS	Comprehensive Adult Student Assessment System
CTE	career and technical education
DOJ	Department of Justice
ESL	English as a second language
GED	General Education Development
Maryland SMS	Maryland Scientific Methods Scale
NAAL	National Assessment of Adult Literacy
OJP	Office of Justice Programs
PLATO	PLATO instructional software package for mathematics, reading, and language
PSE	postsecondary education
RCT	randomized controlled trial
RD	regression discontinuity
SCA	Second Chance Act of 2007 (Public Law 110-199)
SVORI	Serious and Violent Offender Reentry Initiative
TABE	Test of Adult Basic Education
TABE D	Test of Adult Basic Education, Difficult Level

TABE M	Test of Adult Basic Education, Medium Level
WWC	U.S. Department of Education's What Works Clearinghouse

CHAPTER ONE

Introduction

On April 9, 2008, the Second Chance Act (Public Law 110-199) (SCA) was signed into law. This important piece of legislation was designed to improve outcomes for individuals who are incarcerated, most of whom will ultimately return to communities upon release. The Second Chance Act's grant programs are funded and administered by the Office of Justice Programs (OJP) within the U.S. Department of Justice (DOJ). In 2010, for the first time under the SCA, funding was set aside for a comprehensive study of correctional education. OJP's Bureau of Justice Assistance (BJA) awarded the RAND Corporation a cooperative agreement to comprehensively examine the current state of correctional education for incarcerated adults and juveniles and where it is headed, which correctional education programs are effective, and how effective programs can be implemented across different settings. One key task in that effort was to comprehensively review the scientific literature and conduct a meta-analysis to synthesize the findings from multiple studies about the effectiveness of correctional education programs in helping to reduce recidivism and improve employment outcomes.

In this report, we examine the evidence about the effectiveness of correctional education for incarcerated adults in the United States. By correctional education, we mean the following:

- adult basic education (ABE): basic skills instruction in arithmetic, reading, writing, and, if needed, English as a second language (ESL)
- adult secondary education (ASE): instruction to complete high school or prepare for a certificate of high school equivalency, such as the General Education Development (GED)
- vocational education or career and technical education (CTE): training in general employment skills and in skills for specific jobs or industries
- postsecondary education (PSE): college-level instruction that enables an individual to earn college credit that may be applied toward a two-year or four-year postsecondary degree.

Although some may consider life skills programs a part of correctional education, our project focuses specifically on the four types of academic and vocational training programs summarized above. We also limit our focus to correctional education programs provided in the institutional setting, as opposed to postrelease or community-based programs. Finally, our focus is on correctional education programs provided at the state level. These foci enable us to address the question of what is known about the effectiveness of correctional education—specifically, academic programs and vocational training programs—for incarcerated adults in U.S. state prisons.

Our analyses will be of special interest to correctional education administrators, corrections officials, and state policymakers who are interested in understanding the role that correctional education plays in the rehabilitation of and facilitation of incarcerated individuals' reentry back into society and who must carefully consider how they will allocate resources in a fiscally constrained environment. Our findings will inform them about whether there is an association between correctional education and recidivism, postrelease employment, and achievement test scores.

In the remainder of this chapter, we first provide an overview of the field of correctional education. Then, as context for our meta-analysis, we summarize previous meta-analyses that have been done on correctional education. We then summarize the study's objectives and scope, discuss the study's limitations, and describe a roadmap for the remaining chapters.

Background

The growth in the prison population for the past 40 years has been well-documented. In 2010, there were 1.6 million state and federal prisoners in the United States, with more than 700,000 incarcerated individuals leaving federal and state prisons each year (Guerino, Harrison, and Sabol, 2012). About half of state prison inmates in 2009 were serving time for violent offenses, and 19 percent, 18 percent, and 9 percent of state prison inmates were serving time for property, drug, and public-order offenses, respectively. An enduring problem facing the broader system of criminal justice is the high rate of recidivism in the United States: Within three years of release, four out of ten U.S. state prisoners will have committed new crimes or violated the terms of their release and be reincarcerated (Pew Center on the States, 2011). Devising programs and strategies to reduce recidivism requires understanding the unique challenges that individuals face upon release as well the current state of programs in place to mitigate such challenges. We describe both in turn as they pertain to correctional educational programs.

Barriers to Reentry for Incarcerated Prisoners and the Potential of Correctional Education Programs to Address Them

Visher and Lattimore's (2007) evaluation of the Serious and Violent Offender Reentry Initiative (SVORI) found that education, job training, and employment were among the commonly cited needs of incarcerated prisoners reintegrating back into society. But it is challenging to prepare individuals with the needed vocational skills and education to be successful in reintegrating. Ex-offenders, on average, are less educated than the general population (MacKenzie, 2008; Tolbert, 2012). Analysis of data from the Bureau of Justice Statistics' (BJS's) Survey of Inmates in State Correctional Facilities and the National Assessment of Adult Literacy (NAAL) showed that 36.6 percent of individuals in state prisons had attained less than a high school education in 2004 compared with 19 percent of the general U.S. population age 16 and over (Crayton and Neusteter, 2008). Because many inmates lack a high school diploma, the GED certificate is an important way for them to complete basic secondary education (Harlow, 2003). In 2004, 32 percent of state prisoners had earned a GED compared with 5 percent of the general population, whereas only 16.5 percent of state prisoners had a high school diploma compared with 26 percent of the general population (Crayton and Neusteter, 2008). With respect to postsecondary education, 51 percent of the general U.S. adult popu-

lation had at least some postsecondary education compared with only 14.4 percent of state prison inmates.

Literacy levels for the prison population also tend to be lower than that of the general U.S. population. The 2003 NAAL assessed the English literacy of a sample of 1,200 inmates (age 16 and over) in state and federal prisons and a sample of 18,000 adults (age 16 and over) living in U.S. households. Individuals were measured on three different literacy scales: prose, document, and quantitative.[1] On average, inmates had lower scores on all three scales than the general U.S. population (Greenberg, Dunleavy, and Kutner 2007). A higher percentage of the prison population had average scores that fell within the basic level[2] for all three measures of literacy compared with the household population. For example, 40 percent of the prison population was at the basic level for prose literacy compared with 29 percent of the household population; 39 percent of the prison population, for quantitative literacy compared with 33 percent of the household population; and 35 percent of the prison population, for document literacy compared with 22 percent of the household population. All these comparisons were statistically significant (Greenberg, Dunleavy, and Kutner, 2007).

In addition to lower levels of educational attainment, the lack of vocational skills and of a steady history of employment (Petersilia, 2003; Western, Kling, and Weiman, 2001) also represents a significant challenge for individuals returning to local communities (Travis, Solomon, and Waul, 2001). Incarceration affects employment and earnings in a number of ways. Using data from the Fragile Families and Child Wellbeing Study, an analysis of the effects of incarceration on the earnings and employment in a sample of poor fathers found that the employment rates of formerly incarcerated men were about 6 percentage points lower than those for a similar group of men who had not been incarcerated (Gellar, Garfinkel, and Western, 2006). Additionally, incarceration was also associated with a 14–26 percent decline in hourly wages. Given the high incarceration rates in the United States and the fact that many offenders cycle in and out of prison, Raphael (2007–08) noted that the dynamics of prison entry and reentry inhibited the accumulation of meaningful sustained employment experience in this population.

Further, the stigma of having a felony conviction on one's record is a key barrier to postrelease employment (Pager, 2003). Holzer, Raphael, and Stoll (2003) conducted a series of surveys of employers in four major U.S. cities and found that employers were much more averse to hiring ex-offenders than in hiring any other disadvantaged group. Willingness to hire ex-offenders was greater for jobs in construction or manufacturing than for those in the retail trade and service sectors; employers' reluctance was greatest for violent offenders than for nonviolent drug offenders.

Pager (2003) conducted an audit survey of approximately 200 employers in Milwaukee and generated four groups of male job applicants who were very similar in educational and work experience credentials but differed by whether they were offenders or nonoffenders and by race. Pager found that black offenders received less than one-seventh the number of offers received by white nonoffenders with comparable skills and experience. Also, black non-

---

[1] *Prose literacy* measures the knowledge and skills needed to search, comprehend, and use information from continuous texts. *Document literacy* measures the knowledge and skills needed to search, comprehend, and use information from noncontinuous texts. *Quantitative literacy* measures the knowledge and skills needed to identify and perform computations using numbers that are embedded in printed materials.

[2] Literacy levels include Below Basic, Basic, Intermediate, and Proficient.

offenders generated fewer than half as many offers as white nonoffenders—14 percent versus 34 percent, respectively. In terms of differences by racial group, 17 percent of white offenders received a job offer compared with only 5 percent of black offenders. Another barrier is that, in many states, employers can be held liable for the criminal actions of their employees (Raphael, 2007–08). Taken together, lower overall educational attainment, lower levels of literacy, and difficulty securing employment upon release underscores the importance of educational programming for this population.

Overview of U.S. Correctional Education

Most state correctional institutions (84 percent) offer some type of correctional education programming (Stephan, 2008, Appendix Table 18). Data from the BJS 2005 Census of State and Federal Correctional Facilities indicate that 66 percent of state correctional facilities offered literacy or 1st–4th grade education programs, 64 percent offered 5th–8th grade education programs, 76 percent offered secondary or GED, 50 percent offered vocational training, 33 percent offered special education, and 33 percent offered college courses (Stephan, 2008).

Although most state prison facilities offer some form of education, participation rates vary and, in fact, have declined somewhat over time. For example, between 1997 and 2004, participation rates in ABE, GED, postsecondary, and vocational training programs all showed a modest decline (Crayton and Neusteter, 2008). In 2004, 52 percent of state prison inmates reported having participated in a correctional education program since admission to a correctional facility (Harlow, 2003). Only 27 percent of state prison inmates reported having participated in vocational training programs; 19 percent reported having participated in secondary education programs (i.e., high school/GED); 2 percent in adult basic education; and 7 percent in adult postsecondary education programs (Crayton and Neusteter, 2008).

Reasons for the low participation rates may include lack of programs or lack of awareness of program opportunities, reduced funding for correctional education programs because of state budget constraints, or competing demands (e.g., when participation is discretionary, an individual might elect to participate in an employment program rather than an education program) (Crayton and Neusteter, 2008; Tolbert, 2012). In addition, states differ as to whether participation in correctional education programs for incarcerated adults is mandatory or voluntary. A survey of state correctional education programs in 2002 conducted by McGlone found that 22 of the 50 states had adopted legislation or implemented policy requiring mandatory education for prisoners. Of those requiring mandatory participation, ten states had achieving a GED as the requirement for program completion (McGlone, 2002).

The administration and delivery of correctional education also differs from state to state. For example, different entities—state departments of corrections, education, public safety, or labor—may be responsible for administering and financing correctional education programs for their prison systems. Some states have their own correctional school district, such as Texas, Florida, and Ohio. Some states may contract with community colleges to provide GED preparation, postsecondary education, or vocational training programs; other states may contract out only some of their programs. In addition, privately operated corrections firms also have responsibility for providing correctional education to adult prisoners. In 2011, approximately 8 percent of the U.S. state prison population was housed in privately operated facilities (Glaze and Parks, 2012).

Previous Meta-Analyses of Correctional Education

Understanding the role that correctional education plays in rehabilitation and reentry back into society is the key goal of our study and meta-analysis. As a backdrop to our study, we first synthesize findings from previous meta-analyses of correctional education programs in the United States. In keeping with our study goals, we discuss only meta-analyses that have an explicit focus on education programs administered primarily to adult offenders in correctional facilities. According to our review, there have been three major published meta-analyses that meet these criteria: Wilson, Gallagher, and MacKenzie (2000); MacKenzie (2006); and Aos, Miller, and Drake (2006).[3] These studies differ in their parameters, methods, and conclusions. We review the findings from each of these meta-analyses in turn, focusing first on a landmark systematic review of correctional education programs conducted by Lipton, Martinson, and Wilks (1975) that set the stage for the current policy discourse and research direction in the field.[4]

Lipton, Martinson, and Wilks (1975)

In 1975, Douglas Lipton, Robert Martinson, and Judith Wilks published a systematic review of 231 studies of prisoner rehabilitation programs spanning the years 1945 to 1967—a review that provided the first major stocktaking of the potential efficacy of correctional education. Commissioned by the New York State Governor's Special Committee on Criminal Offenders, this seminal review was developed in response to the lack of evidence about whether the array of programs and reform efforts in place at the time were successfully preparing prisoners for reintegration into their communities. For studies to be included in their review, Lipton and his colleagues required that studies use a treatment and comparison group design, with the treatment group composed of program participants and the comparison group composed of nonparticipants. To determine whether different types of programs were working, they tallied the findings from individual studies—those that favored the treatment group, those that favored the comparison group, and those with no discernible difference between the treatment and comparison group—and drew conclusions based on the frequency of statistically significant relationships.

Within their sample of 231 programs, Lipton and his team identified a subset of "skill-development programs," which consisted of academic and/or vocational training. They summarized comparisons of program participants and nonparticipants in studies that used recidivism and employment as outcomes. Across eight studies that assessed recidivism, three showed significantly lower rates of recidivism among program participants, and one showed significantly higher rates of recidivism among program participants. The other four studies showed

---

[3] The studies included in these meta-analyses are largely based on studies of correctional education programs in the United States. However, a handful of international studies are also included.

[4] Since the publication of the landmark Lipton, Martinson, and Wilks study, there have been other systematic reviews of adult correctional education that do not apply meta-analytic methods (e.g., Gaes, 2008), and there have been meta-analyses of correctional education programs administered to juvenile offender populations (e.g., Lipsey, 2009). With the exception of the Lipton, Martinson, and Wilks study, which is important to acknowledge because of its seminal role in the field, we discuss only meta-analyses of adult correctional education programs, because their methods, findings, and conclusions are most relevant for providing context to our study. Additionally, readers should note that we are aware of two dissertations (Chappell, 2003; Wells, 2000) that have used meta-analytic techniques to assess the relationship between correctional education and recidivism. We do not review their analyses in depth here, but their findings, by and large, accord with those of Wilson, Gallagher, and MacKenzie (2000); MacKenzie (2006); and Aos, Miller, and Drake (2006).

no differences between the treatment and comparison groups. In two studies that examined employment as an outcome, offenders who participated in vocational training programs fared worse than nonparticipants after being released. Overall, their review found no conclusive evidence that correctional education was beneficial and found that, in some cases, it might even be harmful.

Lipton's systematic review is notable, in part, because it set the tone for future research and policy discourse in the field. In 1974, one year before the release of the study, Robert Martinson, the study's second author, published a preview of the findings in a commentary "What Works?—Questions and Answers About Prison Reform" in *The Public Interest*. In it, Martinson wrote: "it can safely be said that they [the studies included in their review] provide us with no clear evidence that education or skill development programs have been successful" (p. 27). Martinson's summation cast doubt on the utility of educational programming within the broader system of corrections and generated the provocative conclusion that "nothing works" in prisoner rehabilitation. Although the "nothing works" tagline was never used in the full empirical report, the tagline from Martinson's commentary became synonymous with the Lipton, Martinson, and Wilks review; as a result, federal- and state-sponsored initiatives to address the needs of prisoners were effectively put on the defensive and in some cases curtailed.

Wilson, Gallagher, and MacKenzie (2000)

The empirical documentation of the Lipton study, along with Martinson's critique, galvanized efforts to improve existing academic and vocational training programs and to develop new methods of educating prisoners. However, it was not until 25 years later, in 2000, that the efficacy of correctional education was revisited through a formal meta-analysis conducted by David Wilson, Catherine Gallagher, and Doris MacKenzie (2000) at the University of Maryland. Their meta-analysis included 33 studies of correctional education programs administered to adults published after 1975—a time period that broadly covered the time since the Lipton study was released.

The Wilson, Gallagher, and MacKenzie study attempted to improve on two limitations of Lipton's work: (1) The Lipton study did not address the magnitude of differences in outcomes between treatment and comparison groups, and (2) the Lipton study did not explicitly account for variation in the quality of the research designs across studies. With respect to the former limitation, Lipton's review simply summed up the number of studies that yielded statistically significant differences between the treatment and comparison groups and based the study's conclusions on the preponderance of effects in one direction or the other; this approach is sometimes referred to as a "vote counting" approach, in which each study gets a vote in the "significant" or the "not significant" column, and the votes are counted (Field, 2005). Unfortunately, this approach essentially obscures the magnitude of the effects across studies. In other words, a large difference favoring the treatment group "counts the same" as a small difference favoring the comparison group.

With respect to the latter limitation, Lipton's review discussed differences in methodological quality, highlighting (where appropriate) studies with carefully or poorly selected comparison groups. However, this variation in research design did not factor into how they tallied statistically significant program effects.

To address these limitations, Wilson and his team used formal meta-analytic techniques, which average findings of multiple studies into a single parameter of program or "treatment

group" efficacy.[5] Additionally, they rated each study using a scale that they and their colleagues at the University of Maryland developed specifically for systematic reviews of correctional programs (Sherman et al., 1997). This scale, referred to as the Maryland Scientific Methods Scale (the Maryland SMS), classifies studies as either experimental or quasi-experimental. Following Shadish, Cook, and Campbell (2002), experimental studies are defined as those that randomly assign participants to treatment and control-group status, whereas quasi-experimental studies are those that employ both a treatment and comparison group, but in which group membership is not randomly assigned.

Among the quasi-experimental studies, the Maryland SMS further classifies them according to the quality of statistical controls they employ. Studies from most to least rigorous are classified as follows: Level 5 indicates a well-executed randomized controlled trial (or RCT); Level 4 indicates a quasi-experimental design with very similar treatment and comparison groups; Level 3 indicates a quasi-experimental design with somewhat dissimilar treatment and comparison groups, but reasonable controls for differences; Level 2 indicates a quasi-experimental design with somewhat dissimilar treatment and comparison groups and with limited and/or no controls for differences; and Level 1 indicates a study with no separate comparison group. Wilson and colleagues included only studies that received at least a Level 2 rating and then used the scale as a control variable to determine whether their findings were dependent on the research designs used by the studies' authors.

Whereas the Lipton study documented mostly mixed results, the Wilson study found that correctional programs were beneficial, by and large. In their meta-analysis, they demonstrated that participation in academic programs—including ABE, GED, and postsecondary education programs—was associated with an average reduction in recidivism of about 11 percentage points. This finding was robust when controlling for ratings on the Maryland SMS. Academic program participation was also associated with a greater likelihood of employment, although they did not quantify the relationship in terms of a percentage increase/decrease in the same way they did for recidivism. Vocational training program participation did not yield a consistent relationship with recidivism but was associated with increased odds of employment. Wilson and his team's findings, based on more recent programs and more rigorous methods of analysis, questioned the Martinson study's claim that "nothing works."[6]

MacKenzie (2006)

A few years later in 2006, Doris MacKenzie, a co-author of the Wilson study, updated their original meta-analysis. In this update, she included a handful of newer studies and limited her sample to only those studies published after 1980. Additionally, she limited her sample of studies to only those receiving a Level 3 or higher rating on the Maryland SMS, thereby eliminating studies from the predecessor meta-analysis with Wilson and Gallagher that had the weakest study designs. In her re-analysis, she again found that academic program participation appeared beneficial: The odds of not recidivating were 16 percent higher among academic program participants than nonparticipants. However, with the new sample parameters in place, she now found that vocational program participation was associated with a reduction in recidi-

[5] Meta-analytic techniques were not yet developed at the time of the Lipton study.

[6] Since the publication of the Lipton study, a number of criminologists and policymakers questioned the claim that "nothing works." However, it was not until the Wilson, Gallagher, and MacKenzie study's meta-analysis that a comprehensive evaluation of the literature was synthesized in a systematic way to directly challenge the conclusion of the Lipton study.

vism: The odds of recidivating were 24 percent lower among vocational program participants than nonparticipants. She did not update the analysis of employment.

Aos, Miller, and Drake (2006)

Also in 2006, Steve Aos, Marna Miller, and Elizabeth Drake of the Washington State Institute for Public Policy conducted a meta-analysis of 571 offender rehabilitation programs for adults and for juveniles, ranging from counseling to boot camps to education. They limited their sample to studies conducted from 1970 onward and, like MacKenzie's meta-analysis published the same year, they included only studies that received at least a Level 3 rating on the Maryland SMS. In analyzing 17 studies of academic education programs and four studies of vocational education programs administered to adults, they found results that largely agreed with MacKenzie's: On average, participants have lower rates of recidivism than their nonparticipant peers. Specifically, they found that academic program participation was associated with a 7 percent reduction in recidivism, and vocational program participation was associated with a 9 percent reduction in recidivism.

In sum, early reviews of correctional education programs administered to adults by Lipton, Martinson, and Wilks (1975) found inconclusive evidence to support their efficacy. The lack of consistent positive effects contributed to the popular belief that "nothing works" in prisoner rehabilitation; however, this conclusion may have been premature, given that appropriate analysis techniques had not been developed. More recent reviews using meta-analysis techniques question the conclusions of the earlier work, finding evidence of a relationship between correctional education program participation before release and lower odds of recidivating after release. However, the most recent meta-analyses (Aos, Miller, and Drake, 2006; MacKenzie, 2006) did not consider employment outcomes; thus, whether program participation is associated with postrelease success in the labor market remains unclear.

Study's Objective and Scope

As with the meta-analyses described above, our study aims to understand whether the body of relevant research to date supports the proposition that correctional education programs can help successfully prepare offenders for community reintegration upon release. Following the lead of Wilson and colleagues, MacKenzie, and Aos and colleagues, we use meta-analytic techniques to synthesize the effects of correctional education programs administered to adults across multiple studies. In doing so, our goal is to build on the contributions of their work, while extending them in a number of key ways, which we describe below.

First, our study examines multiple outcomes: recidivism, employment, and achievement test scores. As with previous syntheses, our focus is largely on recidivism, because it is the outcome most often used in the literature, and the ability to avoid recidivism is arguably an important marker of successful rehabilitation. However, we also examine whether participating in a correctional education program is associated with an increase in labor force participation and whether participating in a correctional education program with a computer-assisted instructional component is associated with gains in achievement test scores.

Acquiring steady employment postrelease has been shown to be an important factor in preventing recidivism among ex-offenders (Laub and Sampson, 2003; Uggen, 2000), and among the civilian population, improving the acquisition of academic skills and concepts is

vital in securing employment (Klerman and Karoly, 1994). In terms of life-course or developmental criminology, an emergent body of research has shown that desistance from deviant behavior in adulthood is largely contingent on the opportunity for individuals to acquire new roles and responsibilities in their immediate social nexus. This life-course approach contends that the acquisition of stable, gainful employment—a productive, socially normative role—redirects behavior and energy toward one's family and community and, consequently, away from crime (Laub and Sampson, 2003; Uggen, 2000). We examine employment outcomes, because many of the programs we reviewed were explicitly geared toward providing inmates with occupational skills that they could use to procure employment following release from prison. With respect to skill development, the most proximal measures of program efficacy are indicators of the inmates' learning that can be attributed directly to the courses taken while incarcerated. Thus, our assessment of three distinctive outcomes—recidivism, employment, and academic achievement—helps to elucidate potential mechanisms through which program participation may help improve the postrelease prospects of those formerly incarcerated.

Another way our study differs from the previous meta-analyses is in how we deal with the underlying studies. One major limitation of the extant research on correctional education is the dearth of studies that used experimental designs, making it difficult to establish a causal relationship between program participation and the outcome of interest. Studies that lack experimental designs are susceptible to selection bias, whereby inmates who elect to participate in educational programs may differ in unmeasured ways from inmates who elect not to participate in those programs. For instance, they may be more motivated, have a stronger internal locus of control, or be more proactive about planning for their postrelease futures. Therefore, differences detected between program participants and nonparticipants in meta-analyses with a large number of nonexperimental studies may reflect pretreatment attributes of the inmates who participated in the studies and not the true effects of the programs themselves. To deal with this potential bias, Wilson and colleagues controlled for each study's Maryland SMS rating in their meta-analysis, and the MacKenzie and Aos analyses reviewed only studies that earned at least a Level 3 rating on the Maryland SMS. In our analysis, we pay special attention to those studies receiving a higher (Level 4 or Level 5) rating. As a result, our study provides the most scientifically defensible evidence of program efficacy to date.

A defining feature of our review is that it is the most comprehensive and most recent to date, including a total of 58 studies of correctional educational programs in the United States (compared with 33 studies reviewed by Wilson and colleagues, 22 reviewed by MacKenzie, and 21 reviewed by Aos and colleagues). Our review also focuses specifically on academic and vocational training programs, whereas some of these other reviews also included life skills training/reentry programs and work placement programs. Before our review, the meta-analysis with the most current coverage was Aos, Miller, and Drake (2006), which included studies published through 2005, whereas our meta-analysis incorporates studies published through December 2011. Although this represents a difference of only a few years, it enables us to include 12 newer studies published between 2006 and 2011.

Finally, we used a rigorous review process with multiple quality control checks (described in detail in the next chapter) to ensure that the data extracted from each study are accurate and in accordance with the methods and approaches typically used in the field. Although details on the data extraction process used in previous meta-analyses are limited, it appears that most of this work was carried by the researchers themselves and/or a small team of graduate students. For our study, we assembled an independent scientific review team comprising

content experts external to RAND who have publication and/or funding track records in the field of correctional education research. Each study included in our meta-analysis was assessed independently by two members of the scientific review team, with each independent evaluation reviewed, edited, and finalized by both a graduate student and a project team member. Given the way we constructed the data extraction and review teams and the multiple stages of extraction and review, we feel that the data used to construct our meta-analysis are the most complete in terms of content and quality.

Study's Limitations

As with all studies, there are some study limitations the reader should keep in mind. The majority of the studies we reviewed focused on the outcome measures of recidivism and employment; a more limited set also examined the relationship between correctional education and academic performance. There are also more proximal outcomes of interest in correctional education, such as program completion, behavior while incarcerated, and progress on individual plans and goals. We were limited in our ability to examine these more proximal outcomes because of the limited number of studies examining these indicators.

The correctional education literature is varied, including studies published in academic journals and in other arenas—what often is referred to as the "grey literature." As detailed in Chapter Two, a strength of our study is the literature review process in which we identified studies done on correctional education programs that were published in the peer-reviewed and grey literature by searching online databases, research institutions and colleges' websites, and dissertation abstracts, and by reaching out to departments of corrections and research units. Although our search of the grey literature was extensive, it was not exhaustive, in that we were unable to contact every department of corrections, for example, to obtain copies of unpublished evaluation reports. Of the grey literature we were able to explore, much of it yielded descriptive studies, and our search did not yield studies with research designs of high enough quality to be included in our meta-analysis. That said, to the extent that we missed some high-quality reports from the grey literature through our search strategy, then this is a potential study limitation.

To provide practitioners with evidence on effective program design and implementation and refinement, we originally sought to identify specific aspects of correctional education programs that show signs of efficacy, such as the type of program (e.g., ABE, GED, postsecondary) or the method of delivery used (e.g., whole class instruction, one-on-one instruction). However, few studies provided sufficient information to allow for complete or consistent coding across program characteristics. Despite the need for this information in the field, our analyses are exploratory in nature and limited in what we are able to discern in terms of elements of effective programs.

Finally, our literature review covers the time period from January 1, 1980, through December 31, 2011. As with any systematic literature review and meta-analysis, one has to define a starting point and a cutoff date for inclusion. Our focus on the past three decades precludes a historic look at how correctional education programs may have evolved in the years immediately following the publication of the Lipton study. Additionally, we are aware of a few studies that were just recently published (after our cutoff date of December 31, 2011); these studies were not eligible for inclusion in our meta-analysis.

Organization of This Report

The remainder of this report is organized as follows. Chapter Two summarizes our study methodology. Chapter Three presents the meta-analytic results for the relationship between correctional education and recidivism and the results of a supplementary cost analysis. Chapter Four presents the meta-analytic results for exploring the relationship between correctional education and employment. In Chapter Five, we present the meta-analytic results for computer-assisted instruction and academic performance. In Chapter Six, we provide our overall summary of our meta-analytic findings and discuss policy implications and directions for future research.

This report contains eight appendixes. Appendixes A, B, and C are included as part of this report. Appendix A includes a list of the document identification parameters and sources. Appendix B includes a list of the scientific review team members. Appendix C includes the diagnostic tests for the meta-analyses.

Appendixes D, E, F, G, and H are standalone appendixes posted on the website along with this report. Appendix D includes the scientific review data abstraction protocol. Appendix E includes the list of studies included in the literature review. Appendixes F, G, and H include summaries of the studies included in the recidivism, employment, and computer-assisted instruction meta-analyses.

CHAPTER TWO

Study Methodology

Introduction

This chapter describes our literature search, screening, and review procedures; our approaches to rating the rigor of each study; and the meta-analytic model used to pool and to synthesize the results of these studies. As described in greater detail in this chapter, the meta-analytic results we present are from a comprehensive literature search for published and unpublished studies released between 1980 and 2011 that examined the relationship between correctional education participation and inmate outcomes. We decided to use 1980 as a starting point to ensure that we captured a large enough sample of studies to conduct a meta-analysis with sufficient statistical power; extending too far back in time risks relying on programs that are outmoded and/or less relevant to the current correctional environment. We focused exclusively on studies published in English of correctional education programs in the United States that included an academic and/or vocational curriculum with a structured instructional component.

Studies were subjected to two rounds of screening, each by two independent screeners, for appropriateness of interventions, outcomes, and research designs. Those that met the screening criteria were reviewed independently and in detail by two Ph.D.-level reviewers. The reviews were then reconciled first by a graduate student and then by a Ph.D.-level member of the research team. Outcome data about recidivism rates, employment, and test scores were abstracted and scaled to allow for synthesis across studies, and the meta-analyses were conducted using random-effects pooling.

As with previous meta-analyses that have examined the effects of correctional education described in the previous chapter (Wilson, Gallagher, and MacKenzie, 2000; MacKenzie, 2006; and Aos, Miller, and Drake, 2006), we evaluated the strength of the causal inferences warranted by each study and used these evidence ratings to test the sensitivity of our results to the rigor of the design of the studies. We rated the evidence from each study according to its ability to establish causal inference, using two separate but substantively similar evidence-rating scales—the Maryland Scientific Methods Scale (SMS) (Sherman et al., 1997), which is familiar to those in the criminal justice community, and the U.S. Department of Education's What Works Clearinghouse (2011) rating scheme, which is familiar to those in the field of education. In the remainder of this chapter, we elaborate in greater detail on each step of our methodological approach.

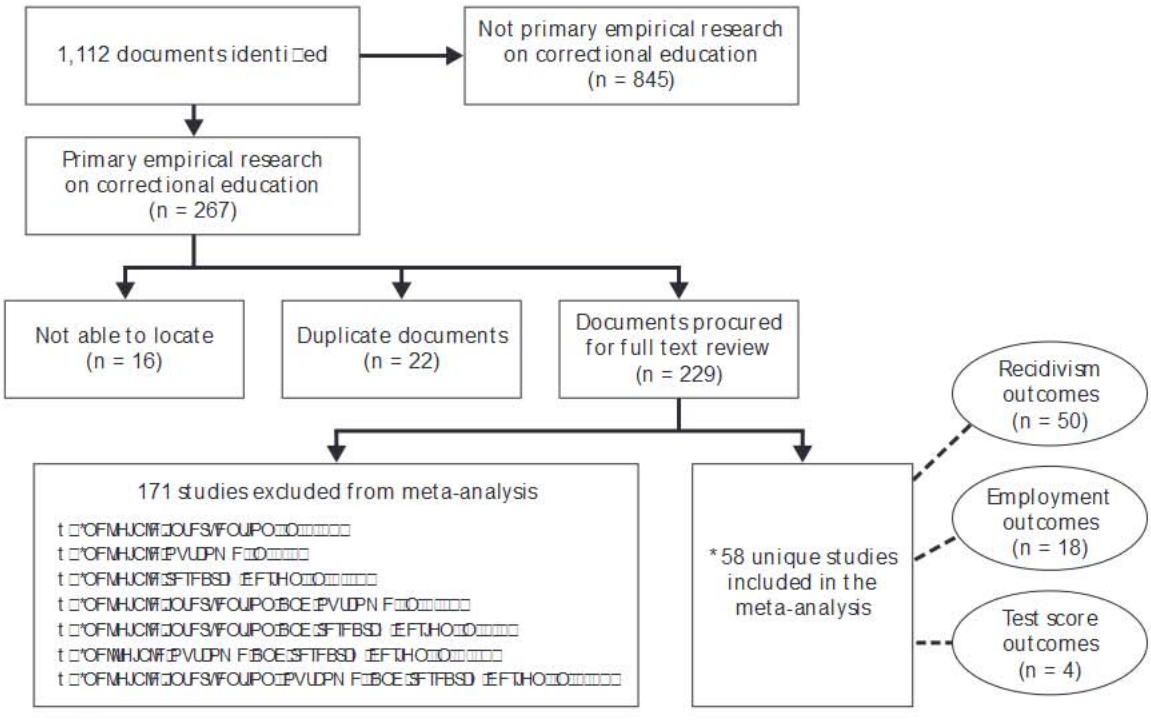
Comprehensive Literature Search

To identify studies for our meta-analysis, we conducted a comprehensive literature search. As part of this search, we first scanned the universe of potential published and unpublished documents to compile all available empirical research studies that examine the effect of correctional education programs on participant outcomes. We then reviewed the documents to determine if they met a set of eligibility criteria that would permit their use in a meta-analysis. A flow chart depicting the steps through which documents were acquired and assessed for eligibility in the meta-analysis is shown in Figure 2.1. We provide details on each of these steps below.

Document Identification

The literature search commenced with an attempt to identify and to locate all possible sources of empirical analyses of correctional education's relationship with inmate outcomes. We employed three methods to identify potential documents, carried out in the following order: a search of relevant research databases, an online repository search, and a bibliography scan. First, we developed a set of search terms (e.g., "correctional education," "prisoner education," "program evaluation") and entered them into search engines of eight research databases widely used by academic researchers. Next, we entered the same set of search terms into online search engines of 11 repositories of criminological research housed at various universities and research organizations. Last, we maintained a record of all major literature reviews and meta-analyses that emerged from the aforementioned database and online repository search. We then searched their bibliographies for potentially relevant citations. A complete list of the search

Figure 2.1
Eligibility Assessment of Potential Documents for Inclusion in the Meta-Analysis



RAND RR266-2.1

terms, research databases, online repositories, and major literature reviews/meta-analyses is included in Appendix A.[1] This document identification stage produced a list of 1,112 citations for documents that could potentially be eligible for inclusion in our meta-analysis.

Eligibility Assessment

Our expansive search strategy yielded a range of documents that were either not focused on correctional education or were not primary empirical studies (e.g., newspaper articles, opinion pieces, literature reviews, workbooks, implementation guides). To eliminate these documents, we trained a team of doctoral students in the Pardee RAND Graduate School (PRGS) on the goals of our review and on how to assess whether the document was a primary empirical study of a correctional education program.[2] To standardize our assessment process, we uploaded the bibliographic reference information for each document into *DistillerSR*, a web-based application designed to facilitate systematic literature reviews. Each reference was assessed independently by two doctoral students within *DistillerSR,* where they had the opportunity to review the document's title, source, and abstract. If they disagreed on whether the document was a primary empirical study related to correctional education, the reference was flagged and a project team member reconciled the discrepancy. If there was not enough information to make a firm assessment, the project team member erred on the side of caution and marked the reference as eligible for the next stage of review.

In this next stage, the list of primary empirical studies of correctional education and the list of references lacking sufficient information to determine if they were primary studies related to correctional education were delivered to RAND's research library staff to retrieve hard copies of the documents. The documents were then uploaded into *DistillerSR*. For this second round of review, two doctoral students independently evaluated the full text of the document. With access to the entire document in addition to the bibliographic reference information, they were able to confirm whether or not it was a primary empirical study of correctional education.

Of the original 1,112 documents identified, 845 were not primary empirical studies of correctional education and 267 were primary empirical studies of correctional education. Of the 267 primary empirical studies, we were unable to locate 16 documents, and an additional 22 documents were determined to be duplicates of other studies. This included either exact duplicates or studies by the same author(s) that were published in different venues but with the same findings and/or analytic samples. In the latter situation, we used the document with the most comprehensive information on the program and the study design. For each of the 229 nonduplicative studies that we were able to obtain, the doctoral students examined its content to determine if it met three criteria necessary for inclusion in our meta-analysis:

---

[1] In addition to our systematic approach to identifying potential documents, a number of researchers and practitioners directly provided us with documents for consideration (most of which had already been identified through our database search strategy). This cooperation was due to the high visibility of the project among members of the Correctional Education Association and the Association of State Correctional Administrators. All documents, regardless of how they were identified, were subjected to the same eligibility assessment procedures.

[2] We define a *primary empirical study* as one in which the authors were directly responsible for the research design, data analysis, and the reporting of the findings.

- The study needed to evaluate an *eligible intervention*.
- The study needed to measure success of the program using an *eligible outcome measure*.
- The study needed to employ an *eligible research design*.

For our study, we define an *eligible intervention* as an educational program administered in a jail or prison in the United States published (or released) between January 1, 1980, and December 31, 2011. We define an educational program as one that includes an academic and/or vocational curriculum taught by an instructor, designed to lead to the attainment of a degree, license, or certification. The program could be part of a larger set of services administered to inmates or it could be a stand-alone program. However, it needed to have an explicit academic or vocational curriculum in place with an instructional component. Therefore, prison work programs and job placement programs lacking a structured training component under the supervision of an instructor were deemed ineligible. Additionally, although the program may include postrelease services, it must be primarily administered while the inmate is held in a correctional setting. Programs administered to parolees were excluded. Instructional programs that did not explicitly address academic or vocational skills—for instance life skills, drug rehabilitation, and anger management programs—were also excluded.

The study needed to measure the effectiveness of the program using an *eligible outcome measure*, which for our meta-analysis include recidivism, employment, and achievement test scores. Initially we kept our parameters broad and considered a range of possible outcomes, such as disciplinary infractions while incarcerated, postrelease educational attainment, wages, and subjective evaluations of program effectiveness. However, a representative meta-analysis requires a moderate number of studies with outcomes measured in a comparable way. Few studies with these other outcomes met this requirement, and so we eventually excluded them from consideration.

For the purposes of our meta-analysis, we consider an *eligible research design* as one in which there is a treatment group composed of inmates who participated in and/or completed the correctional education program under consideration and a comparison group composed of inmates who did not. Comparison groups that deviated from this definition—such as comparison groups composed of nonincarcerated participants or comparison groups who received a different correctional education intervention from the one under consideration—were not eligible. In reporting the findings, the authors of the study needed to include sufficient statistical detail on both the treatment and the comparison groups to meet this eligibility criterion.[3]

As with the initial review of the bibliographic reference information, if the two doctoral students reviewing the full text of the document disagreed on whether the document met any of these three criteria, the document was flagged and a project team member reconciled the discrepancy. Of the 229 nonduplicative studies that we were able to obtain, 58 studies had an eligible intervention, an eligible outcome measure, and an eligible research design—and, thus, were eligible for inclusion into our meta-analysis. Of these 58 studies, 50 studies used recidivism as an outcome variable, 18 studies used employment as an outcome variable, and four studies used achievement test scores as an outcome variable. All four of the studies that used achievement test scores as the outcome variable evaluated the effects of computer-assisted instruction. Therefore, although our analyses of recidivism and employment outcomes look at

[3] If they reported means for the treatment and comparison group, we required that they also provide sample sizes. If they reported a regression coefficient, we required that that they also provide a standard error.

a broad range of correctional education programs, our analysis of achievement test scores is solely focused on programs with computer-assisted instruction. Hence, we refer to our analysis of test scores as our computer-assisted instruction meta-analysis. Bibliographic citations for all 229 nonduplicative, locatable primary empirical studies of correctional education and their status with respect to the three eligibility criteria are reported in Appendix E. Those 58 studies deemed eligible for meta-analysis were then subjected to a formal scientific review, described in detail in the next section.

Scientific Review

Independent Reviews by the Scientific Review Team

Once the studies had been screened for eligibility, those deemed eligible were reviewed in greater detail by two researchers who independently extracted information about the interventions, outcomes, and participants in each study. To undertake these detailed reviews, we appointed a scientific review team made up of ten faculty members from various academic departments across the country who possessed not only methodological expertise in quantitative social science research but also substantive expertise in correctional education, criminal justice, and/or social services for at-risk populations. A list of the scientific review members, their educational credentials, and their current positions is shown in Appendix B.

To guide extraction of the data from the individual studies, we designed a scientific review protocol. This protocol was developed with close attention to the review procedures used in the U.S. Department of Education's What Works Clearinghouse (2011), as well as the procedures used in the University of Maryland's "Preventing Crime" report (Sherman et al., 1997). The resulting protocol, which is displayed in Appendix D, included four worksheets. The first or *main worksheet* contains 44 questions, most of which are multiple choice. These questions focus largely on the characteristics of the program being evaluated, as well as on the study's setting, design, and publication venue. The scientific review team helped guide the selection of intervention characteristics so that our analysis would be as useful as possible to policymakers and practitioners. The *outcomes worksheet* asks for information about the outcome variables in the study. The *baseline characteristics* worksheet captures descriptive information about the study participants. Finally, the *reviewer log* asks about the reviewers' overall impressions of the strengths, weaknesses, and implications of the study.

The scientific review process commenced with two full days of training for the team members on how to use the scientific review protocol to record relevant data from the studies. Following the training, reviewers independently completed two practice protocols, and we provided the team with detailed feedback about response patterns and guidance to encourage standardized answers. To further encourage consistency among reviewers, we provided a written manual that further clarified the intent of each question in the protocol.

After two scientific review team members independently reviewed each eligible study, the main worksheets of the two independent reviews were merged into a single, reconciled review. A project team member then examined each review, referring back to the material in the original document to reconcile items on which the two independent reviewers provided substantively different responses. Another project team reconciled the outcomes and baseline characteristics worksheets, in all cases consulting each original study to ensure correct data extraction. As a final precautionary measure, the dataset of extracted, reconciled outcome and

baseline characteristics information was checked twice against the main text of the studies for data recording accuracy.[4]

Defining Treatment and Comparison Groups

As described above, our meta-analysis is founded on the aggregation of studies that include both a treatment group consisting of inmates who participated in and/or completed a correctional education program and a comparison group consisting of inmates who did not participate in and/or complete the correctional education program. Most studies compared outcomes between these two mutually exclusive groups to test the hypothesis that exposure to correctional education improved outcomes. In some cases, the study included more-refined groups based on treatment dosage and program completion. For example Cronin's study of GED programs in Missouri (2011) identified four groups of inmates: (1) inmates who came to prison without a GED and did not make any progress; (2) inmates who came to prison without a GED, made progress toward obtaining GED, but did not earn a GED; (3) inmates who earned their GED in prison, and (4) inmates who came to prison with a GED or more. In this instance, we constructed our treatment and comparison groups as conservatively as possible following an intent-to-treat approach. In an intent-to-treat approach, every subject who was assigned to the treatment group is analyzed on the outcome of interest as a member of the treatment group, regardless of whether they received the full dosage of the treatment through completion. In accord with this approach, we coded groups 2 and 3 in Cronin (2011) as the treatment group and group 1 as the comparison group. Thus, our analysis reflects all inmates without any exposure to a GED program (comparison group) to inmates who were exposed to any amount of correctional education while incarcerated, regardless of whether they completed the program (treatment group).

Rating the Quality of the Research Design

The quality of any meta-analysis depends on the quality of studies it includes (LeLorier et al., 1997; Slavin, 1984). One particular concern in social science research—and by extension, in social science meta-analysis—is that effects attributed to program participation in the original studies may actually be driven by the types of individuals who elect to participate in the program rather than by the causal effect of the program itself. This problem is typically referred to as selection bias. To minimize concerns about selection bias, some researchers advocate strict restrictions on the quality of studies included in the meta-analysis (Slavin, 1984), such as the exclusion of all studies that are not RCTs.

Often considered the gold standard in social science research, RCTs are desirable because the random assignment of research participants to treatment and control groups renders the two groups identical in expectation at the time of assignment (Shadish, Cook, and Campbell, 2002), allowing us to reasonably infer that any average differences in their outcomes were attributable to the intervention (Myers and Dynarski, 2003; Shadish, Cook, and Campbell, 2002). In practice, of course, treatment and comparison groups cannot be infinitely large,

[4] In a number of cases, the data provided in the article were insufficient for direct use in a meta-analysis and needed to be recalculated or recalibrated so that they could be consistently input into the analysis as odds ratios. For example, some articles provided means without standard errors, or regression coefficients without the total number in the study. In these cases, we performed our own calculations. Hence, some of our reported estimates for each article differ somewhat from what was included in the original publication.

so there is the potential for treatment and comparison groups to differ as a result of random variation. In addition, RCTs sometimes suffer from attrition after the point of randomization, which can potentially introduce systematic differences between the two groups. Despite these limitations, RCTs offer a strong defense against selection bias, because the treatment assignment process is, by definition, independent of the characteristics of the participants.

Other rigorous comparison-group designs, such as regression discontinuity designs and instrumental variables analysis, attempt to minimize selection bias in nonrandomized studies by capitalizing on arguably random processes, but in doing so, they must satisfy a larger set of assumptions to nullify the threat of selection bias (Angrist and Pischke, 2009; Murnane and Willett, 2011; Schochet et al., 2010). Still other designs attempt to mitigate selection bias comparing the treatment group to a non–randomly assigned comparison group that is *observably* similar. Some studies achieve this through matching or weighting the comparison group so that it is similar to the treatment group on a number of possibly confounding characteristics. When the number of characteristics to be used in the weighting or matching is large, balance can sometimes be achieved by using these characteristics to estimate the probability of receiving the treatment and matching treated to comparison cases based on these fitted probabilities, or propensity scores (McCaffrey, Ridgeway, and Morral, 2004; Rosenbaum and Rubin, 1983; Rubin, 1997). Matching or weighting on observed characteristics helps ensure that the observed characteristics are not responsible for any apparent treatment effects, but it leaves open the possibility that unmeasured differences may be driving such effects. Moreover, because researchers rarely have comprehensive measures of all the group differences—such as motivation, perseverance, time orientation, or locus of control—that may drive selection into the groups and also be associated with outcomes, matching and weighting studies remain vulnerable to selection bias.

Similarly, studies that use covariate adjustment—that is, that statistically control for possible confounding characteristics through multivariate regression—are also vulnerable to biases from unobserved variables. In comparison to matching or weighting studies, those that use regression controls may also be more vulnerable to misspecification of the functional form of the relationship between variables—that is, to incorrectly assuming particular linear or curvilinear relationships (Ho et al., 2007). Although some studies have found that matching and weighting perform little better than covariate adjustment, given the same variables (Cook et al., 2009), there remains a preference in the field for balanced treatment and comparison groups over reliance on statistical controls to adjust for differences between dissimilar groups (What Works Clearinghouse, 2011).

Given the centrality of selection bias as a threat to causal inference in the literature on social and educational interventions, we rated the quality of evidence in each reviewed study based on how well the study's design mitigated this threat. Specifically, we sought to classify the rigor of the eligible studies using evidence ratings that focused on the warranted strength of the causal inference and could be well-understood by both the criminal justice and education communities.

As noted in Chapter One, we chose the Maryland Scientific Methods Scale, which was developed for the 1997 *Preventing Crime* report published by University of Maryland researchers (Farrington et al., 2002; Sherman et al., 1997). The Maryland SMS rates studies on a five-point scale, where Level 5 is the most rigorous, indicating a well-executed randomized controlled trial with low attrition; Level 4 is a quasi-experimental design with very similar treatment and comparison groups; Level 3 is a quasi-experimental design with somewhat dis-

similar treatment and comparison groups but reasonable controls for differences; Level 2 is a quasi-experimental design with substantial baseline differences between the treatment and comparison groups that may not be well controlled for; and Level 1 is a study with no separate comparison group that does not receive the treatment. As noted in Chapter One, the Wilson, Gallagher, and MacKenzie (2000) meta-analysis was restricted to studies rated Level 2 or higher on the Maryland SMS, and the later meta-analyses by MacKenzie (2006) and by Aos and colleagues (2006) included only studies rated Level 3 and higher.

For communicating results in a way that would be easily understood by the education community, we also used the U.S. Department of Education's What Works Clearinghouse rating scheme—herein referred to simply as the WWC rating scheme for ease of expression. The WWC rating scheme has only three categories: Meets Standards, Meets Standards with Reservations, and Does Not Meet Standards. A study that *Meets Standards* on the WWC rating must be a randomized, controlled trial with low levels of overall and differential attrition or it must use a well-executed regression discontinuity or single-case design. An RCT that exceeds the attrition threshold (described further below) is reviewed as a quasi-experimental design.[5] A study *Meets Standards with Reservations* if it is a quasi-experimental design in which the treatment and comparison groups are observably very similar at the point of analysis. This means that all observed baseline characteristics for the treatment and comparison groups are within 0.25 of a standard deviation of each other and that there are statistical controls for any differences greater than 0.05 of a standard deviation. A study in which the treatment and comparison groups are not within 0.25 of a standard deviation of each other on all observed baseline characteristics and lack statistical controls for any differences greater than 0.05 of a standard deviation *Does Not Meet Standards.*

Operational Use of the Maryland SMS and WWC Rating Scheme

A useful feature of the Maryland SMS and WWC rating scheme is that their two highest evidence categories correspond very closely. Of the two, however, the WWC rating scheme is more specific than the Maryland SMS about precise cutoffs regarding baseline equivalence and attrition.

Baseline equivalence refers to the degree to which the treatment and comparison groups are similar at the beginning of the study in terms of characteristics known to influence the outcome. If a study uses random assignment to assign participants to the treatment and comparison groups, then baseline equivalence is assumed by both the Maryland SMS and WWC rating scheme. This is because random assignment ensures that self-selection is not driving membership in the treatment or comparison group at the point of assignment. The groups differ in expectation only by their assignment status, which is random by design. Of course, as noted above, differences may result simply by accident, especially when the groups are small. For this reason and because it improves the precision of the treatment effect estimate, researchers often adjust for any *observed* baseline differences even in the case of random assignment.

---

[5] The WWC rating scheme maintains newer sets of standards for two other research designs that can warrant causal inference. These are regression discontinuity designs, in which assignment to the treatment or comparison group depends on falling immediately on either side of a numeric threshold, such as a test score cutoff (Schochet et al., 2010), and single-case designs, which lack an untreated comparison group but in which causality is established by repeatedly introducing and withdrawing the treatment from the participants in one of several patterns (Kratochwill et al., 2010). Although the former are increasingly popular in policy analyses (Angrist and Pischke, 2009), and the latter are popular in special education research, no eligible studies with either of these designs were uncovered in our comprehensive literature search.

However, neither the WWC rating scheme nor the Maryland SMS requires adjustment for baseline differences in cases of random assignment.

In studies that do not have random assignment, baseline equivalence is established by demonstrating that the treatment and comparison groups are observably similar on key variables that may be related to both treatment status and the outcome variable, since baseline differences between groups could bias the treatment effect estimates, as noted above. For example, if inmates who enroll in correctional education have lower baseline education levels than those who do not, then any differences in the two groups' outcomes could be due to their prior education levels (and associated aptitude or motivation levels) as much as to the effect of the treatment program. Both the Maryland SMS and WWC rating scheme are based largely on the strength of evidence about baseline equivalence, with randomized designs receiving the highest ratings.

Attrition rates refer to the percentage of participants whose outcomes are lost to the study for any number of reasons, such as inability to collect follow-up data on the inmate, transfer of the inmate to a different correctional facility, loss of follow-up data, and so forth. Importantly, attrition is not the same as noncompletion of a program or intervention among those whose outcomes are observed. Noncompleters who drop out of an intervention program are viewed simply as noncompliant treatment recipients, and they are defined as part of the treatment group within our intent-to-treat framework.

The WWC rating scheme is concerned with two types of attrition: overall attrition and differential attrition. Both may undermine the advantages of random assignment by introducing self-selection into the sample for which outcomes are observed. Overall attrition is simply the total share of baseline participants lost to the study; differential attrition is the percentage-point difference in the attrition rates of the treatment and comparison groups. Because the concerns about attrition pertain to disruption of random assignment advantages, we follow the WWC rating scheme in applying attrition calculations only to studies that begin with a randomized design. Randomized trials with low overall and differential attrition meet the highest standards on the WWC rating scheme, and we apply this standard to the Maryland SMS as well to meet its highest category. Studies that do not begin with a randomized design do not need to meet an attrition threshold, but they are also ineligible for the highest ratings on either the Maryland SMS or the WWC rating scheme. These studies need only establish strong evidence of baseline equivalence to meet the second-highest tiers of evidence on both scales.

To summarize, studies with high rates of attrition and/or that lack baseline equivalence may yield biased results. Because the WWC rating scheme offers clear guidelines to establish specific numeric thresholds for these validity threats, we apply those thresholds to both scales. Our operational definitions of each scale are presented in Table 2.1.

Because the Maryland SMS' and WWC rating scheme's evidence standards are quite similar, we operationalize the strongest evidence categories identically across the two scales. To receive the highest evidence rating on each scale, a study must meet the liberal standard for low overall and differential attrition to earn a *Meets Standards* rating on the WWC rating scheme and a Level 5 on the Maryland SMS.[6] WWC has not published the precise formula for its attrition thresholds, so the formulae we used are extrapolated from the attrition macro in the

[6] The WWC maintains both a liberal and a conservative threshold for attrition (see Appendix A in What Works Clearinghouse, 2011). Both thresholds are designed to keep attrition-related bias within 0.05 of a standard deviation of the outcome measure, but the liberal threshold is based on less pessimistic assumptions about selective attrition. We chose to apply the

Table 2.1
Operational Definitions of Evidence Rating Categories in the What Works Clearinghouse Rating Scheme and the Maryland Scientific Methods Scale

What Works Clearinghouse Rating Scheme	Maryland Scientific Methods Scale	Joint Operational Definition
Meets standards	5	Randomized controlled trial with attrition below the liberal WWC threshold
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
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 for matching, and no statistical controls for baseline differences between treatment and comparison groups
	1	No separate comparison group

template provided to WWC reviewers, with confirmatory reference to the attrition threshold graphics in version 2.1 of the WWC *Procedures and Standards Handbook* (What Works Clearinghouse, 2011).[7] The formulae we used are as follows:

Low attrition:
Differential attrition rate ≤ 0.129 – (0.192 * Overall attrition rate)

High attrition:
Differential attrition rate > 0.129 – (0.192 * Overall attrition rate)

where 0.129 represents the y-intercept of the attrition threshold (i.e., the acceptable level of differential attrition, defined as 12.9 percentage points difference between the treatment and comparison groups), and –0.192 represents the slope, or the difference in the differential attrition level associated with a unit difference in the overall attrition rate.

In each formula, as noted, the *overall attrition rate* is the pooled sample of study participants included in the final analysis divided by the pooled sample at the point of randomization. The *differential attrition rate* is the absolute value of the attrition rate for the treatment group minus the corresponding rate for the comparison group. We also operationalize a Maryland SMS Level 4 study and a study that Meets WWC Standards with Reservations identically across the two scales. Studies at this level are quasi-experimental designs in which the

liberal threshold because there are so few RCTs in correctional education research. For example, of the four RCTs identified for the meta-analyses, three met the liberal threshold for attrition, but only one would have met the conservative threshold.

[7] The graphic depicting the thresholds changed slightly from version 2.0 to version 2.1 of the *Procedures and Standards Handbook*, with no corresponding change in the text or definition of the thresholds, and inquiries to the WWC for the precise equation were unsuccessful. For increased precision, we ultimately used a formula extrapolated from the macros in the 2010 study review guide (the data-extraction tool provided to reviewers), although our ratings of the studies were not sensitive to small variations in the threshold formula.

treatment and comparison groups are observably very similar, primarily because of deliberate matching or weighting of the comparison group to the characteristics of the treatment group. The result should be that treatment and comparison group members differ by no more than 0.05 standard deviation units on three baseline dimensions that are known to be related to recidivism outcomes and that are relevant to an educational intervention: namely, *age*, *prior offenses*, and *baseline educational level*. (This also requires that standard deviations of the baseline characteristics be reported in the studies.) Moreover, we specify that the recidivism and employment studies must take into account the *length of time between release and data collection*, since inmates released for longer periods will have more time to recidivate and/or to find work. They can do this by observing everyone for a certain time period (e.g., one year postrelease) or through survival analysis methods that adjust for duration of release. Because correctional facilities are typically gender-segregated, gender is an unlikely source of selection bias in this context. Therefore, matching on or controlling for gender does not affect a study's evidence rating in our analysis.

It is important to note that in requiring baseline equivalence on only four variables, we depart slightly from the WWC guidelines, which require that *all* observed baseline characteristics—whether 1 or 50, for example—fall within 0.25 standard deviations of each other for the treatment and comparison group and that differences of more than 0.05 of a standard deviation be held constant statistically. The reason for this departure is that very few studies in our sample provided adequate information about the distribution of baseline characteristics for us to run these calculations, but a number of studies described matching procedures to ensure close balance of the treatment and comparison groups on particular variables. Also, we do not require matching on all observable variables, because this penalizes studies that report larger numbers of variables—with a large enough set of variables, we would expect some differences by chance alone, and this chance would be greater in smaller studies, even when the studies were otherwise equivalent. Instead, we set a consistent expectation of matching or achieving strong similarity on the three variables known to be strong predictors of postrelease outcomes and on the time period that the individuals were observed postrelease.

Studies below Level 4 on the Maryland SMS are classified as "Does Not Meet Standards" by the WWC rating scheme, because these categories do not require strong similarity between the treatment and comparison groups. We operationalize a Level 3 study on the Maryland SMS to be one that includes statistical controls for at least one of the aforementioned key baseline differences between groups and/or includes matching or weighting on one or two of these variables.

We classify Level 2 studies as those that include nonrandomly assigned treatment and comparison groups but do not include any statistical controls or adjustments for differences between groups. Finally, we classify Level 1 studies as those that lack a comparison group consisting of inmates who did not receive the treatment.[8]

Although we classify all studies included in our meta-analysis on both the WWC rating scheme and the Maryland SMS, we organize most of our analyses around the Maryland SMS because of its granularity in classifying studies—allowing us to make comparisons at more refined levels of study design rigor. As Wilson and colleagues (2000) did, we restrict our meta-analysis to studies rated a Level 2 or higher on the Maryland SMS, effectively limiting it to

[8] Note that Level 1 would not include single-case design studies. Had we encountered such studies in our search and screening processes, they would have been rated separately according to the WWC standards for single-case designs.

studies that include a distinct comparison group that did not receive the treatment. However, we focus particularly on the Level 4 and Level 5 studies, which are the least vulnerable to selection bias. We consider the results from these higher-quality studies to be the most robust for use and application in the field. However, the inclusion of the lower-quality studies in some specifications ensures that we are also making use of the findings of a broad set of studies of a range of program types and models undertaken during the last 32 years.

Description of the Data

As shown in Figure 2.1, we determined that 58 studies were eligible for inclusion into our meta-analysis. For analytic purposes, however, our unit of analysis is the effect size (k) and not the individual study (n). An effect size is the statistic reported in the study that indicates the magnitude of the difference on the outcome of interest between a treatment group and a comparison group. Across the 58 studies, we were able to extract a total of 102 effect sizes. The number of effect sizes exceeds the number of studies, because a study could contain multiple treatment and comparison groups and thus multiple comparisons. For example, a study making a single comparison of recidivism rates between a treatment group receiving GED coursework and a comparison group receiving no GED coursework would contribute only one effect size to our meta-analysis. However, a study comparing the recidivism rates of two treatment groups—one receiving GED coursework and one receiving vocational certification training—with the recidivism rate of a comparison group receiving no form of correctional education would contribute two effect sizes to our meta-analysis.

Our recidivism analysis is based on 71 effect sizes from 50 studies, our employment analysis is based on 22 effect sizes from 18 studies, and our test score analysis is based on nine effect sizes from four studies. Table 2.2 shows the distribution of studies and effect sizes according to their rating on the Maryland SMS and the WWC rating scheme.[9] The majority of studies are of recidivism and employment and the majority of effect sizes come from Level 2 and Level 3 studies on the Maryland SMS and Do Not Meet Standards according to the WWC rating scheme—suggesting that, on average, the field of correctional education research is limited in its ability to assess whether correctional programs yield a causal effect on recidivism and employment. Therefore, in our analysis, we focus in where possible on those studies that receive a Level 4 or Level 5 rating.

Analytic Approach

We conducted our meta-analysis using a random-effects approach. Random effects meta-analysis is appropriate when effect sizes are heterogeneous. This might occur when the individual studies are not sampled from the same population; this can be conceptualized as there being a "super-population" of all potential respondents, which contains an array of subpopulations, and each study randomly samples from one of these subpopulations. In addition, dif-

[9] Note that in Table 2.2, the distribution of studies (n) across the Maryland SMS ratings for the recidivism analysis sums to 51 and not 50. This is because one study (Piehl, 1995) contributed two effect sizes that had different Maryland SMS ratings and therefore appears in two separate rows.

Table 2.2
Distribution of Studies and Effect Sizes, by Rating Categories in the What Works Clearinghouse Rating Scheme and the Maryland Scientific Methods Scale

What Works Clearinghouse Rating Scheme	Maryland Scientific Methods Scale	Recidivism Analysis		Employment Analysis		Test Score Analysis	
		n	k	n	k	n	k
Meets standards	5	2	2	0	0	2	4
Meets standards with reservations	4	5	7	1	1	1	3
	3	20	29	9	11	0	0
Does not meet standards	2	24	33	8	10	1	2
	1	na	na	na	na	na	na
Total sample		51	71	18	22	4	9

NOTES: n is the number of studies, k is the number of effect size estimates, and na is not applicable. Studies receiving a Level 1 on the MD Scale do not include any type of comparison group; therefore, there was no way to calculate an effect size estimate. They were excluded from our analysis by design. The n column in the Recidivism Analysis column does not sum to 50 because one study (Piehl, 1995) contributes two effect sizes at different rating levels.

ferences in treatment protocols or contexts might also introduce heterogeneity. For our meta-analysis, we consider the super-population to be all inmates in correctional facilities in the United States between 1980 and 2011, and the subpopulations might be minimum-security inmates in California in 1985; medium-security inmates in Connecticut in 2003, etc. Rather than assuming that each study has randomly sampled from the super-population, we consider that each study has sampled from one of the subpopulations. Hence, there is substantial heterogeneity in the effect size estimates across the different subpopulations.

Random-effects models are an appropriate technique for meta-analysis when there is substantial heterogeneity in effect size estimates across the different subpopulations, as is the case in our review of correctional education programs.[10] We use a DerSimonian-Laird estimator to pool results across the multiple effect sizes. This estimator weights each study's effect size estimate by the precision (e.g., standard error), and the heterogeneity of effect sizes (e.g., gives greater weight to those studies that are closer to the mean), and then produces a pooled effect size and standard error. This pooled effect size in our meta-analysis provides an estimate of the relationship between participation in correctional education and our three outcomes across the population of eligible studies. Because of the nested nature of our data (e.g., multiple effect sizes within the same study), the assumption of independent observations is violated, which may result in artificially narrow standard errors. To assess this, as a sensitivity analysis we computed robust standard errors using robust hierarchical meta-analysis (Hedges, Tippton, and Johnson, 2010).[11]

[10] Random-effects models was also the estimation method used in three major meta-analyses published to date (Wilson, Gallagher, and MacKenzie, 2000; MacKenzie, 2006; and Aos, Miller, and Drake, 2006).

[11] We computed robust standard errors for meta-regression using the ROBUMETA command available in Stata (Hedberg, 2011). This was necessary only for our analysis of recidivism, as there was not sufficient nesting in the pool of eligible studies of employment or test scores to permit this computation. The results were not contingent on the method for estimating the standard errors; tests of significance reflect unadjusted standard errors.

One limitation of systematic reviews is that studies that fail to produce statistically significant results have a more difficult time getting published in journals—leading to publication bias or "the file drawer problem" (i.e., studies that find no program effects remain in file drawers and are not widely distributed). This publication bias may skew the findings in favor of successful programs. We attempted to limit the threat of publication bias by searching an array of sources in the literature to procure official program evaluation reports not published in journals, working papers, research briefs, theses, and dissertations.

To assess whether our results are contingent on the studies that we were able to procure, we perform two diagnostic tests. Our first diagnostic test assesses whether studies with positive results have a higher probability of publication—that is, whether we can find evidence of publication bias. Large studies, which have more power and smaller standard errors, will have a greater chance than small studies of obtaining a statistically significant result, if the population effect size is equal in those studies. If there is no publication bias, the average effect size estimate of the smaller studies in our pool of eligible studies should be the same as the average effect size estimate of the larger studies in our pool of eligible studies. If publication bias is having an effect, then small studies that do not obtain statistically significant results should have a lower chance of being published. This can be depicted visually in a "funnel plot" and formally tested using either a parametric test (Egger et al., 1997) or a non-parametric test (Begg, 1994).

A second diagnostic test we perform is a "leave-one-out" analysis. There is a risk that one large study with an extreme result may bias the results of the analysis. To ensure that this is not the case, we run "leave-one-out" analysis, in which the data are re-analyzed leaving out studies one at a time, until all studies have been excluded individually. We then ensure that the substantive conclusions are unchanged, regardless of which studies are included or excluded. The results from these diagnostic tests and their implications for interpreting the main analytical findings are shown and described in Appendix C. In short, there is some evidence of publication bias in the body of studies on recidivism, but this bias is small and unlikely to substantively change the results of our main findings.

CHAPTER THREE

The Relationship Between Correctional Education and Recidivism

Introduction

This chapter presents the results from our meta-analysis where recidivism is the outcome. We first describe how we defined and measured recidivism across the 50 eligible studies and then pool all 71 effect size estimates from the 50 studies together to provide an aggregate estimate of the relationship between participation in correctional education and recidivism. Next, we examine the relationship when restricting only to studies with the most rigorous research designs. We then use previously published national estimates of recidivism to help interpret the magnitude of this relationship. We also explore whether the relationship between correctional education and recidivism varies by the type of program and instructional delivery method used. We conclude with a straightforward cost analysis that compares the cost of correctional education to the cost of reincarceration.

Measuring Recidivism

Recidivism was measured in many ways across the 50 eligible studies along three dimensions: the *definition of recidivism* used by the researcher, the *time period* between release from prison and when recidivism is recorded for study participants, and the *statistical metric* used by the researcher to report the degree of recidivism experienced by the treatment and comparison group members. We describe each of these dimensions below in turn.

- *Definition of recidivism*. Recidivism is defined a number of ways, including reoffending, rearrest, reconviction, reincarceration, technical parole violation, and successful completion of parole. In our pool of 50 studies that had recidivism outcomes, the majority used reincarceration as the outcome measure (n = 34).
- *Time period*. Studies varied in the time period through which they followed the study participants after release from prison, which represents their time "at risk for recidivism." Studies ranged from examining a cohort of former inmates in the community for six months since release from prison to following them for over ten years since release from prison. The most frequently used time periods in the 50 eligible studies were one year (n = 13) and three years (n = 10).
- *Statistical metric*. Forty-two of the studies reported the percentage in treatment and comparison group that recidivated and seven of the studies reported regression coefficients along with standard errors to express the magnitude of the difference in recidivism between the treatment and the comparison groups. One study (Piehl, 1995) contributed

effect sizes reported two different ways—one based on a percentage comparison between the treatment and comparison group and the other based on a regression coefficient.

When there were multiple outcomes and reporting methods used, we gave preference to reincarceration (as this represents the modal definition of recidivism), recidivism within one year of release or as close as possible to one year (as this represents the modal time period used by the authors of the studies), and regression coefficients (as this represents the best attempt by the authors of the studies to reduce potential sources of bias). When these were unavailable, we used whatever definition, time period, or statistical metric reported by the author so that we could be as inclusive as possible. As such, our recidivism measure comprises a range of slightly different measures, and thus should not be interpreted in terms of the individual measures that make it up.[1] Details on how each of the 50 studies defined and operationalized recidivism, as well as specific information on the individual programs being studied, the research design used in the study, the WWC's and the Maryland SMS' ratings of the study's research design, and the rates of recidivism recorded for the treatment and comparison group are shown in Appendix F.

We transformed all 71 effect size estimates from the 50 studies into 71 odds ratios.[2] Recall that the number of effect sizes exceeds the number of studies, because a study could contain multiple treatment and comparison groups and thus multiple comparisons. For our purposes, the odds ratio is calculated as the odds of recidivating among treatment group members divided by the odds of recidivating among comparison group members. Odds ratios greater than 1 indicate that the treatment group had a higher rate of recidivism, and odds ratios less than 1 indicate that the comparison group had a higher rate of recidivism. An odds ratio of 1 indicates that there is no difference between the treatment group and the comparison group.[3] These 71 odds ratios form the data points on which the random-effects regression is estimated.

---

[1] Our aggregation of multiple types of recidivism and time periods is based on the assumption that the estimated effect of correctional education is not contingent on the measurement strategy or specification used by the researcher. We tested this assumption by sampling studies that reported the effects of correctional education on recidivism using different definitions and time periods. We found that the effect of correctional education did not differ across the definition of recidivism (e.g., reincarceration, rearrest, parole failure) or time period used (e.g., six months since release from prison, one year since release from prison, ten years since release from prison). This gives us confidence that the findings from our meta-analysis are robust and apply to a range of postrelease settings, circumstances, and outcomes.

[2] We use log odds ratios in producing our analysis, because they have a symmetrical distribution and an associated standard error. We convert these log odds ratios into odds ratios before presenting and interpreting the relationships, as the log odds ratio has no straightforward, intuitive interpretation.

[3] For example, in Torre and Fine's (1997) study of female inmates who enrolled in a postsecondary education program in New York state, the authors found that 7.7 percent of the treatment group returned to prison within three years of release and that 29.9 percent of the comparison group returned to prison within three years of release. The odds associated with a 7.7 percentage are 0.077 / (1 – 0.77) = 0.083; in other words, the odds of a treatment group member recidivating are 0.083 to 1. The odds associated with a 29.9 percentage are 0.299 / (1 – 0.299) = 0.43; in other words, the odds of a comparison group member recidivating are 0.43 to 1. The associated odds ratio for this effect size estimate is 5.12 (0.083 ÷ 0.43 = 0.19) and indicates that the odds of recidivating among treatment group members is 0.19 times than the odds of recidivating among comparison group members. The actual odds ratio for Torre and Fine (1997) as shown in Figure 2.2 is 5.11; this is the reciprocal of the result we give, as we give the odds of recidivating, whereas that study presents the odds of not recidivating. The two analyses are equivalent.

Results: Estimates of the Relationship Between Correctional Education and Recidivism

The Overall Relationship Between Correctional Education and Recidivism

To assess the relationship between correctional education and recidivism, we first graphed the odds ratios for each of the 71 effect size estimates in Figure 3.1 using a forest plot. Each row in the plot corresponds to an effect size, labeled on the left with the corresponding first author of the study and the year of publication. Studies with multiple effect sizes are listed multiple times with a capital letter to differentiate among them. The black box represents the effect size estimate for the study, and the "whiskers" extend to the range of 95 percent confidence intervals.[4] The size of the box is proportional to the weight that is assigned to that effect size. Weight is determined by sample size, and in the case of a random effects regression such as this, the weight is determined by the difference between the estimate of the effect in that study and the overall aggregated effect across studies. A very large study, such as Allen's (2006) study of over 16,000 inmates participating in academic programs in 15 states, is highly weighted and is represented with a large box.

The box and whiskers for each effect size are plotted in relation to the dashed line down the center of the graph, which indicates an odds ratio of 1. Effect sizes to the right of this line indicate that the treatment group had a higher odds of recidivating, and effect sizes to the left of this line indicate that the comparison group had a higher odds of recidivating. If the whiskers for the corresponding box do not cross this dashed line, then the study yielded a significant difference between the treatment and comparison group for that particular effect size at the conventional level of $p < 0.05$. Conversely, if the whiskers for the corresponding box cross this dashed line, then there is no significant difference detected between the treatment and comparison group for that particular effect size at the conventional level of $p < 0.05$.

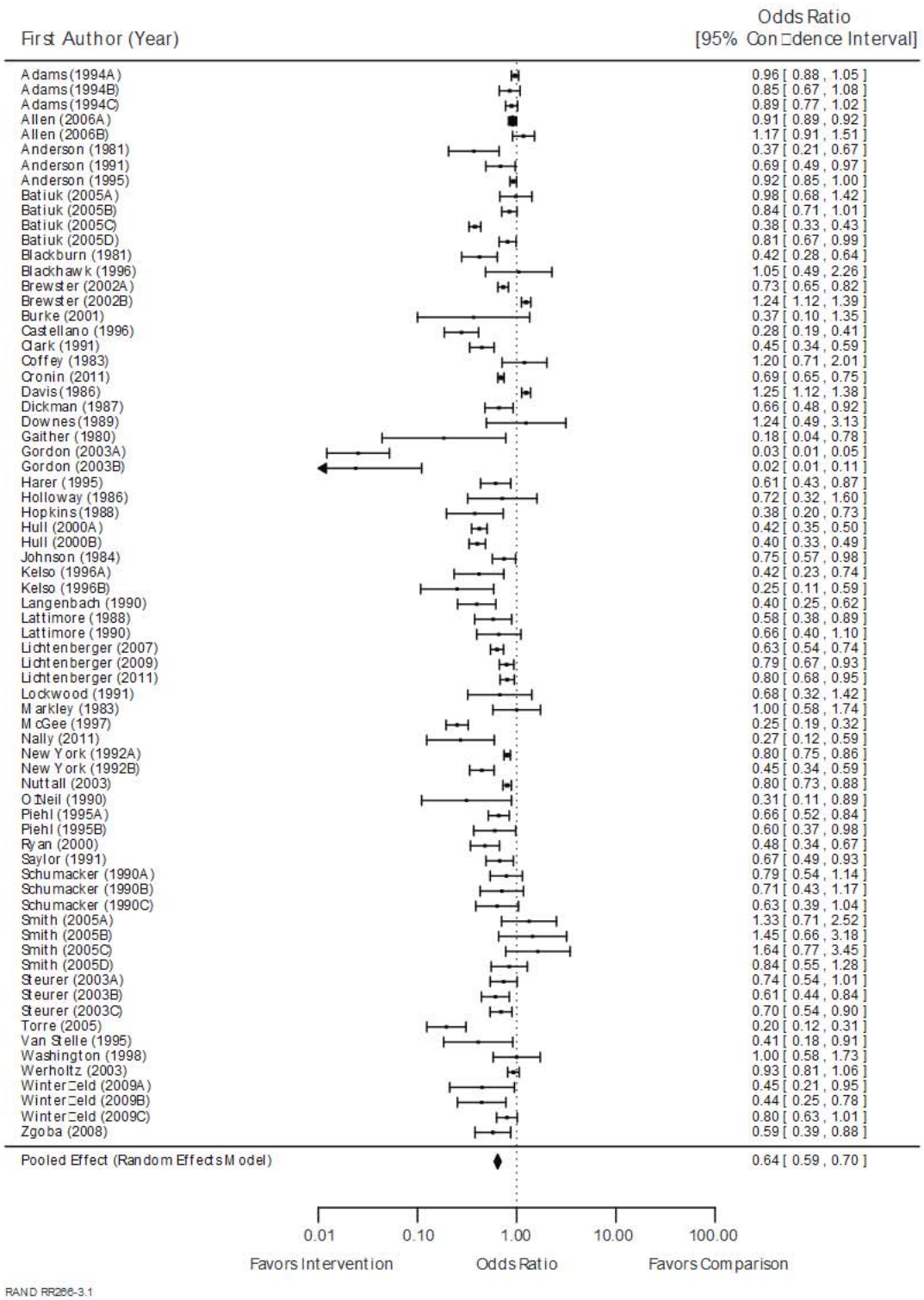
As can be seen by the patterning of boxes and whiskers in this figure, the majority of studies report that the odds of recidivism are lower in the treatment group, with one study (Gordon and Weldon, 2003) finding substantially lower odds of recidivism among treatment group members. A small number of studies find lower odds of recidivism in the comparison group, but these do not generally achieve statistical significance, as evidenced by the fact that the corresponding whisker crosses the solid black line. The very last row displays the overall odds ratio for all 50 studies with 71 effect size estimates pooled together. The position of this overall odds ratio is indicated across the rest of the studies by the diamond at the bottom of the graph. The overall odds ratio is 0.64 ($p < 0.05$, 95 percent confidence interval = 0.59 to 0.70), indicating that across 32 years of empirical studies on the effects of correctional education with analyses ranging in methodological quality and rigor, *on average, the odds of recidivating among inmates receiving correctional education are 64 percent of the odds of recidivating among inmates not receiving correctional education.*

The Relationship Between Correctional Education and Recidivism in Studies with High-Quality Research Designs

As described above, many studies have limitations in their research design that preclude them from ruling out selection bias as an explanation for the observed differences between the treat-

[4] Note that the left whisker for Gordone (2003b) is an arrow. This is to signify that the confidence interval for this effect size extends beyond the scale of the figure.

Figure 3.1
Odds Ratios for Each of the 71 Effect Size Estimates



RAND RR266-3.1

ment and comparison groups. Therefore, although we find across the full sample of studies that participation in correctional education is associated with a reduction in the odds of recidivism following release, we also examine whether this pattern is maintained when we restrict our sample to studies with the strongest and most scientifically defensible research designs. To this end, we recalculated the odds ratio for studies that fall at different levels of the Maryland SMS. We first show the odds ratio for those reaching a Level 5—the highest level of methodological rigor. We then recalculated the odds ratio for studies reaching both Level 4 and Level 5. From here, we stepwise recalculated the odds ratio to incrementally include each of the lower levels of the Maryland SMS. The odds ratios and their corresponding confidence intervals are shown in Table 3.1. The bottom row in Table 3.1 shows the odds and ratio and confidence interval for all studies meeting a Maryland SMS Level 2 and above, which includes all 50 studies and 71 effect size estimates and represents the overall aggregated odds ratio as originally reported in Figure 3.1.

We focus our attention on studies that receive a Level 4 or Level 5 rating on the Maryland SMS, as they are the most methodologically rigorous and provide the best estimate of the causal relationship between correctional education and recidivism. Level 5 consists of experimental studies that employ randomized control designs, and those in our systematic review that are eligible for the recidivism meta-analysis include two studies with two corresponding effect sizes. Both studies evaluate the Sandhills Vocational Delivery System Experiment in North Carolina (Lattimore, Witte, and Baker, 1988; 1990). The odds ratio for these two studies is 0.61 ($p < 0.05$, 95 percent confidence interval = 0.44 to 0.85), indicating that the odds of recidivating among treatment group members in these experimental studies are 61 percent of the odds of recidivating among comparison group members.

Although Level 5 on the Maryland SMS reflects the most stringent research design, the estimate is less informative, because it is based on only one program and, hence, is restricted in its broader applicability to the array of correctional education programs in operation. To incorporate a broader range of programs while maintaining a high degree of methodological rigor, we focus on Level 4 and Level 5 studies combined. Level 4 consists of quasi-experimental studies where the treatment and control group are reasonably matched on a number of key observable characteristics. Among those eligible for the recidivism meta-analysis, five studies receive a Level 4 rating: Harer's (1995) study of federal prison education programs (including Adult Basic Education, GED, postsecondary education including college courses and vocational training), Langenbach et al.'s (1990) study of televised postsecondary instruc-

Table 3.1
Estimates of the Effect of Correctional Education Participation on the Odds of Recidivating, by Levels of the Maryland Scientific Methods Scale

Maryland Scientific Methods Scale	Odds Ratio	95% Confidence Interval	n	k
Level 5	0.61*	0.44 to 0.85	2	2
Levels 4 and 5	0.57*	0.47 to 0.70	7	9
Levels 3, 4, and 5	0.68*	0.60 to 0.78	27	38
Levels 2, 3, 4, and 5 (total sample)	0.64*	0.59 to 0.70	50	71

*$p < 0.05$.

NOTE: n is the number of studies and k is the number of effect size estimates.

tion in Oklahoma state prisons, Nally et al.'s (2011) study of Indiana Department of Corrections' education programs (including Adult Basic Education, GED, postsecondary education including college courses and vocational training), Saylor and Gaes' (1996) study of the Post-Release Employment Project vocational training program administered in federal prisons, and Winterfield et al.'s (2009) study of prison postsecondary education in Indiana, Massachusetts, and New Mexico.

When we combine these five Level 4 studies with the two Level 5 studies, our aggregated odds ratio is 0.57 ($p < 0.05$, 95 percent confidence interval = 0.47 to 0.70), indicating that the odds of recidivating among treatment group members in the most-rigorous quasi-experimental studies are 43 percent lower than the odds of recidivating among comparison group members. That we obtain odds ratios that are of similar magnitude when restricting our analysis to the studies with the most rigorous research design suggests that the overall effect observed among our full sample of 50 studies is not driven by lower-level studies that are potentially subject to selection bias.

Despite the robustness of our findings across levels of the Maryland SMS, we cannot say definitively that the similarity of estimates among the lower-level and higher-level studies means that the programs in each group are equally effective. For example, it is possible that the estimates for the lower-level studies are inflated by selection bias and that the estimates for the higher-level studies generalize only to particular types of higher-quality programs. Yet a closer examination of these studies shows that programs in the higher-level and lower-level studies are similar on most attributes we recorded.[5] This suggests that the programs are not drastically different in the two groups of studies and that the effect estimates in the lower-level studies are relatively unbiased.

Interpreting the Relationship Between Correctional Education and Recidivism

Because the odds of an outcome—in our case, recidivating—can be a less-intuitive metric to grasp, we applied two other metrics to aid in interpretation: the *risk difference* and the *number needed to treat*. The *risk difference* is the absolute reduction in recidivism rates between those who received correctional education and those who did not. The *number needed to treat* indicates the predicted number of inmates who need to receive correctional education to prevent one additional inmate from recidivating. These two metrics require an estimated rate of recidivism in the population upon which to calibrate their calculations.[6] We used recidivism rates from two studies to translate our odds ratio into a risk difference and number needed to treat:

---

[5] In analyses not shown, we find no statistically significant differences in program characteristics at the 5 percent level between higher-level and lower-level studies in terms of the type of instructor (college, certified, corrections officer, outside employee, volunteer), the type of instruction (whole class, small group, one-on-one), the academic or vocational emphasis of the program, and the presence of postrelease supports. In addition, we find that the studies are similarly likely to have missing data on these variables and on the jurisdiction of the facility (federal, state, local). However, the two statistically significant differences that we do find between higher-level and lower-level studies are in the share of programs in federal prisons (i.e., two programs, accounting for 44 percent of the effect estimates, in higher-level studies, versus none in the lower-level studies), and in the security level of the prisons. (In the higher-level studies, we find 44 percent of effects have missing security-level data, and none come from programs in maximum-security facilities. In the lower-level studies, we find 76 percent with missing security-level data and the remainder of programs in a roughly equal combination of minimum, medium, and maximum security facilities.) It therefore remains possible that the effects from the higher-level and lower-level studies reflect the effects of different kinds of programs or contexts.

[6] To take an extreme example, if only 1 percent of inmates recidivated, and education programs prevented all recidivism, we would need to treat 100 inmates to expect to stop one inmate from recidivating. At the other extreme, if the recidivism

rearrest rates and reincarceration rates from Langan and Levin's (2002) study for the Bureau of Justice Statistics and reincarceration rates from a more recent study conducted by the Pew Charitable Trusts (Pew Center on the States, 2011). We base our calculations on our odds ratio for those studies meeting a Level 4 or Level 5 rating on the Maryland SMS, as these represent our best estimate of the causal effect of correctional education on recidivism using an array of programs. We present these additional interpretative metrics in Table 3.2.

Recidivism rates from the aforementioned published studies indicate that between 43.3 percent and 51.8 percent of former prisoners were reincarcerated within three years of release, and two-thirds were rearrested within three years of release. If we apply the recidivism rates estimated by Langan and Levin (2002) for the Bureau of Justice Statistics, we find that *correctional education would be expected to reduce three year rearrest and reincarceration rates by 13.2 and 13.8 percentage points, respectively.* According to these estimates, eight inmates would need to receive correctional education to prevent one additional inmate from being rearrested within three years of release, and seven inmates would need to receive correctional education to prevent one additional inmate from returning to prison within three years. The magnitude of these effects is similar when considering more recent national level recidivism estimates by the Pew Charitable Trusts (Pew Center on the States, 2011): Correctional education would be expected to reduce three-year reincarceration rates by 12.9 percentage points and eight inmates would need to receive correctional education to prevent one additional inmate from returning to prison within three years.

Role of Program Type and Instructional Delivery Method

Though the effect size estimates shown in Figure 3.1 favor the intervention in the majority of cases, resulting in a positive average effect across studies, it is important to note that the estimates are heterogeneous. That is, some are more positive than others, and a few are null or

Table 3.2
Risk Difference and Number Needed to Treat Based on Different Recidivism Base Rates

Recidivism Base Rate Source	Recidivism Base Rate Definition	Recidivism Base Rate	Estimated Recidivism Rate for Correctional Education Participants	Risk Difference	Number Needed to Treat
P. A. Langan and D. J. Levin, Recidivism of Prisoners Released in 1994, NCJ 193427, 2002	Rearrest within 3 years of release	67.5%	54.3%	13.2%	8
P. A. Langan and D. J. Levin, Recidivism of Prisoners Released in 1994, NCJ 193427, 2002	Reincarceration within 3 years of release	51.8%	38.0%	13.8%	7
Pew Center on the States, State of Recidivism: The Revolving Door of American Prisons, Washington, D.C.: Pew Charitable Trusts, 2011	Reincarceration within 3 years of release	43.3%	30.4%	12.9%	8

NOTE: Risk Difference and Number Needed to Treat estimates are based on the odds of recidivating among correctional education participants in seven studies that meet a Level 4 or Level 5 rating on the Maryland SMS.

rate were 100 percent, we would need to treat only one inmate to have the same expected reduction in recidivism. Therefore, even though the effects of the treatment are the same, the cost-effectiveness is dependent on the rate of recidivism.

negative. This heterogeneity may be driven by a variety of factors, including variation in the program features, their contexts, and/or how they are implemented. To help states and localities develop effective programs, it is important to use what we know about the programs to interpret the sources of this variation.

A core focus of policymakers and practitioners in the field of correctional education is developing programs that are designed and delivered in a manner that can yield the most benefit. To help inform decisions about program attributes, we sought to identify whether certain characteristics of programs were more or less associated with reductions in recidivism. When abstracting information on the individual programs into the review protocol (shown in Appendix D), the scientific review team members identified the type of program examined (e.g., GED preparation, vocational training) and the instructional delivery method used (e.g., whole class instruction, one-on-one class instruction).[7] We use this information to recalculate our odds ratios for programs with these different characteristics. Because of the small number of studies that provided information on their programs, we based these analyses on all studies eligible for the recidivism analysis (i.e., those with at least a Level 2 rating on the Maryland SMS) to provide sufficient sample sizes for analysis. Because of the small sizes and potential bias (stemming, perhaps, from researchers who provide more information on program characteristics because they are likely more closely connected with the program), we urge readers to interpret these findings with caution.

Program Type

We calculate odds ratios for four types of correctional education programs: ABE programs, high school/GED programs, postsecondary education programs, and vocational education programs. The odds ratios are presented in Table 3.3. A limitation in interpreting these odds ratios is that studies differed in the precision in which they classified their programs. For example, some studies focused exclusively on a particular vocational program in which participants were exposed only to an occupationally focused curriculum with complementary job training

Table 3.3
Estimates of the Effect of Correctional Education Participation on the Odds of Recidivating, by Program Type

Program Type	Odds Ratio	95% Confidence Interval	n	k
Adult basic education	0.67*	0.57 to 0.79	13	19
High school/GED	0.70*	0.64 to 0.77	22	28
Postsecondary education	0.49*	0.39 to 0.60	19	24
Vocational education	0.64*	0.58 to 0.72	34	42

*$p < 0.05$.

NOTE: n is the number of studies and k is the number of effect size estimates.

[7] As shown in the review protocol in Appendix D, the scientific review team abstracted a range of details about the programs in each study. Ideally, we would like to report on all program characteristics collected to provide a more comprehensive understanding of what is most effective in correctional education. However, few studies provided sufficient information to allow for complete or consistent coding across these characteristics. We present only the analyses for program type and instructional delivery method if they had a minimum of four effect size estimates.

and counseling, whereas other studies focused on broader correctional education programs that included vocational courses taken alongside a set of academic courses. A study of the latter type would therefore be included in the vocational education category as well as in one of the other program categories. Consequently, the independent effects of the vocational and academic components would remain inseparable because the studies do not generally disaggregate the effects of each component or report on individual-level dosage and outcomes in a way that would allow our analysis to disaggregate the effects. Because of the overlap in curricular exposure and the lack of specificity in dosage, the odds ratios for the different program types *should not be compared directly with one another*. In other words, we cannot say with certainty that the programs grouped in each category are pure examples of a given program type (e.g., adult basic education or postsecondary education). Rather, they are programs that include at least some components of that program type.

The results in Table 3.3 suggest that participation in a correctional education program—regardless of the type of program—is associated with a reduction in recidivism. All four of the odds ratios for program type are less than 1 and are statistically significant at $p < 0.05$. Although different programs serve inmates with different needs and skill sets—e.g., postsecondary education programs are typically administered to the most academically advanced inmates and ABE programs are typically administered to inmates with low levels of academic attainment—the findings here suggest that correctional education may be an effective way to prevent recidivism for prisoners across the spectrum of ability and academic preparedness.

It is worth noting that the U.S. Department of Justice (Harlow, 2003) reports that approximately 68 percent of inmates in state prisons lack a high school diploma. Therefore, high school/GED programs would be the most relevant and common approach to educating the majority of prisoners. In our meta-analysis, we were able to identify 28 effect size estimates from 22 studies of high school/GED programs. The associated odds ratio for these programs is 0.70 ($p < 0.05$, 95 percent confidence interval = 0.64 to 0.77), indicating that the odds of recidivating among inmates participating in high school/GED programs are 70 percent of the odds of recidivating among similar inmates not participating in such programs.

Instructional Delivery Method

We next calculate odds ratios for seven instructional delivery approaches. The odds ratios corresponding to these methods are presented in Table 3.4. Similar to the analysis of program type, these methods are not mutually exclusive. For example, some programs use whole class instruction or one-on-one instruction and provide a postrelease component. Hence, the odds ratios should not be compared directly with one another, and thus *it is not appropriate to conclude that certain delivery methods are more or less effective than others*. Five of the delivery methods yield statistically significant odds ratios: programs that use whole class instruction, programs with courses taught by college instructors, programs with courses taught by correctional employees, programs with courses taught by instructors external to the correctional facility, and programs that have a postrelease component. The other two methods—one-on-one instruction and classes taught by certified teachers—do not appear to result in a significant reduction in recidivism among treatment group members. One-on-one instruction is likely administered to inmates with the greatest developmental needs, and so the lack of a difference between the comparison and treatment group can potentially be considered a sign of progress (assuming that the comparison group comprises largely inmates without develop-

Table 3.4
Estimates of the Effect of Correctional Education Participation on the Odds of Recidivating, by Instructional Delivery Method

Instructional Delivery Method	Odds Ratio	95% Confidence Interval	n	k
Whole class instruction	0.71*	0.55 to 0.93	10	13
One-on-one instruction	0.98	0.80 to 1.21	5	8
Class taught by certified teacher	1.14	0.82 to 1.57	1	4
Class taught by college teacher	0.44*	0.33 to 0.59	11	12
Class taught by correctional employee	0.65*	0.50 to 0.85	9	14
Class taught by outside employee	0.54*	0.42 to 0.70	12	17
Program has postrelease services	0.43*	0.30 to 0.62	7	13

*$p < 0.05$.

NOTE: n is the number of studies and k is the number of effect size estimates.

mental needs). Although we do not find a statistically significant effect for programs that use certified teachers, this is based on a single study.[8]

A common thread among three of the five statistically significant instructional delivery methods—programs with courses taught by college instructors, programs with courses taught by instructors external to the correctional facility, and programs that have a postrelease component—is that they connect inmates both directly and indirectly with the outside community. College instructors and instructors external to the facility can potentially infuse the program with approaches, exercises, and standards being used in more traditional instructional settings. Additionally, these instructors provide inmates with direct, on-going contact with those in the outside community. Programs with a postrelease component provide continuity in support that can assist inmates as they continue on in education and/or enter the workforce in the months immediately after they are released. Although we are limited in our ability to classify programs and to establish causality, the findings here provide suggestive evidence that correctional education may be most effective in preventing recidivism when the program connects inmates with the community outside the correctional facility.

Comparison of the Costs of Correctional Education and Reincarceration Costs

To place our meta-analytic findings into context, we undertook a straightforward cost analysis using estimates of the costs of correctional education and of reincarceration.[9] The cost analysis is done for a three-year window after release from prison.

[8] As context, it is worth nothing that within the field of education research, the evidence is mixed as to whether teacher certification matters for student achievement (Seftor and Mayer, 2003).

[9] Although our meta-analysis incorporated a range of indicators to construct our measure of recidivism (e.g., reincarceration, rearrest, parole revocation rates), here we are able to base our cost analysis on estimates of cost for three-year reincarceration rates.

To determine the average cost of providing education to inmates, the average rate of reincarceration, and the average cost of reincarceration (see Table 3.5), we obtained the following three inputs. First, we required an estimate of the cost per year per inmate for correctional education. We used data from Bazos and Hausman (2004) who calculated the average cost of correctional education programs per inmate participant using information from The Three States Study, which assessed the relationship between correctional programs and recidivism in Maryland, Minnesota, and Ohio for approximately 3,170 inmates (Steurer, Smith, and Tracy, 2003). We also used data from the 2007 Corrections Compendium Survey Update on Inmate Education Programs (Hill, 2008). These two sources estimated that the average annual cost of correctional education programs per inmate participant was $1,400 and $1,744, respectively.

Second, the reincarceration rate affects the cost-effectiveness of the intervention: The higher the reincarceration rate, the greater the potential cost savings. We used the three-year reincarceration rate estimates presented in Table 3.2 for correctional education participants and nonparticipants. Specifically, we used the most conservative reincarceration rate estimates based on the Pew Charitable Trust's most recent national estimate of reincarceration based on 41 states: 43.3 percent for individuals who *did not* receive correctional education, and 30.4 percent for who those *did*—a risk difference of 12.9 percentage points as estimated from our meta-analysis (Pew Center on the States, 2011).

Third, we used data on the average annual cost per inmate of incarceration from the Bureau of Justice Statistics' (Kyckelhahn, 2012) analysis of state corrections' expenditures[10] and the Vera Institute of Justice study on the price of prisons (Henrichon and Delaney, 2012), which collected cost data from 40 states using a survey; these two studies estimated the average annual cost per inmate to be $28,323 and $31,286, respectively. Assuming a mean incarceration length of stay of 2.4 years (Pastore and Maguire, 2002), we calculated the average incarceration costs as between $67,975 and $75,086, respectively, based on the two studies.

Table 3.5
Inputs into the Cost Analysis

Input	Lower-Bound Scenario	Upper-Bound Scenario
Cost of Providing Education to Inmates		
Average annual cost of education per inmate	$1,400	$1,744
Average Rate of Reincarceration		
Three-year reincarceration rate	Nonparticipants: 43.3% Participants: 30.4%	
Average Cost of Reincarceration		
Average annual cost of incarceration per inmate	$28,323	$31,286
Average incarceration cost per inmate assuming an average length of stay of 2.4 years	$67,975	$75,086

[10] Expenditure data were extracted from the U.S. Census Bureau.

We applied these three inputs to a hypothetical pool of 100 inmates to calculate cost savings estimates (presented in Table 3.6). We estimated that 43.3 percent of individuals who did not receive correctional education would be reincarcerated within three years, leading to reincarceration costs of between $2.94 million and $3.25 million (Table 3.6).[11] If correctional education were offered to these inmates, our estimates suggest that the reincarceration rate might drop to 30.4 percent giving rise to incarceration costs of between $2.07 million and $2.28 million—a difference of $0.87 million (using lower-bound estimates) or $0.97 million (using upper-bound estimates). Thus, the costs of providing education to this group of 100 inmates would range from $140,000 to $174,000. This translates as a per inmate cost ranging from $1,400 to $1,744, suggesting that providing correctional education is cost-effective compared with the cost of reincarceration.

Another way to look at it is to calculate the break-even point—that is, the risk difference in reincarceration rate required for the cost of education to be equal to the cost of incarceration (as shown in the equation below).

$$\text{Risk difference required for cost effectiveness} = \frac{\text{cost of education}}{\text{cost of incarceration}}$$

For a correctional education program to be cost-effective (from a fiscal/correctional budgetary standpoint alone), it would need to reduce the three-year reincarceration rate by between 1.9 percentage points (using the lower-bound estimate of the cost of education and the upper-bound estimate of the cost of incarceration) and 2.6 percentage points (using the lower-bound estimate of the cost of incarceration and the upper-bound estimate of the cost of education). In fact, our meta-analytic findings indicate that participation in correctional education programs

Table 3.6
Cost Analysis Results

	Lower-Bound Estimate	Upper-Bound Estimate
Reincarceration costs for participants not participating in correctional education[a]	$2.94 million	$3.25 million
Reincarceration costs for those participating in correctional education[b]	$2.07 million	$2.28 million
Difference in costs between the two groups	$0.87 million	$0.97 million
Cost of providing correctional education to the 100 inmates	$140,000	$174,400
Cost of providing correctional education per inmate	$1,400	$1,744

[a] Assumes that 43.3 percent of correctional education nonparticipants would be reincarcerated within three years.

[b] Assumes that 30.4 percent of correctional education participants would be reincarcerated within three years.

[11] The correct numbers to use here are the marginal costs, not average costs, but marginal costs are not readily available. For educational programs, marginal costs are probably similar to average costs. For incarceration, marginal costs may be somewhat lower than average costs.

is associated with a 13 percentage point reduction in the risk of reincarceration three years following release.

A full analysis of the benefits and costs of correctional education was beyond the study's scope. Besides accounting for the direct costs to a prison system, such an analysis would also need to account for other costs, such as the financial and emotional costs to victims of crime and to the criminal justice system as a whole, which could be much more substantial than our estimates above. Also, because few studies have investigated the effect of education for more than three years, we assumed that the effect of correctional education programs after three years is equal to zero. However, these programs may have a "protective effect," diminishing the odds of reincarceration for some years after release.

For ease of calculation, we assumed that the effects of program participation were uniform across different types of crimes. However, a richer treatment of the issue would consider the possibility of heterogeneous effects across crimes and across individuals with different profiles. (It may be that education works better for people who have a lower-than-average tendency to recidivate to begin with.)

In addition, a full benefit and cost analysis would need to account for the dynamics of how people move in and out of prison over their lifetimes. Most studies look at the reduction in reincarceration rates over a short period of time (e.g., one-year). However, there is a lack of good data on lifetime reincarceration rates. Last, a full benefit and cost analysis would need to factor in the costs associated with crime-causing activity that does not result in incarceration. In the late 1970s, RAND conducted prisoner surveys in Texas, Michigan, and California. Using self-reported data, RAND found that the median number of crimes (excluding all drug crimes) reported by prisoners in the year before their incarceration was 15.[12] Data from more recent studies on self-reported criminal activity have yielded similar results (DiIulio and Piehl, 1991; Levitt, 1996). Our analyses did not take into account the number and types of crimes prevented by providing correctional education to prisoners.

Summary

When examining 71 effect size estimates from 50 studies of correctional education programs spanning 32 years of research with analyses ranging in methodological quality and rigor, the majority of studies we identified showed lower rates of recidivism among inmates receiving correctional education than among inmates who did not receive correctional education. To provide the best estimate of the causal relationship between correctional education and recidivism, we examined nine effect size estimates from seven studies that received a Level 4 or Level 5 rating on the Maryland SMS (i.e., the most rigorous research designs) and found that the odds of recidivating among treatment group members are 43 percent lower than the odds of recidivating among comparison group members. When applying these estimated odds to the most recently reported national rates of reincarceration (43.3 percent within three years of release), correctional education would reduce reincarceration rates by 12.9 percentage points on average, although effectiveness does appear to differ by program.

Our findings complement those detected in the most recent meta-analyses published by Wilson, Gallagher, and MacKenzie (2000); Aos, Miller, and Drake (2006); and MacKenzie

[12] That is, the median number of crimes committed that were not caught or prosecuted.

(2006)—all of which document that correctional education participants have lower rates of recidivism than nonparticipants. Unfortunately, all of these studies disaggregate their point estimates differently and do not use the same metric to report their findings. Hence, it is not possible to directly compare the size of the estimates across studies. However, that four independently conducted meta-analyses with different methods and criteria yield consistent results lends weight to the proposition that correctional education can reduce the likelihood that inmates will return to crime upon release.

To place our meta-analytic findings into context, we undertook a cost analysis using estimates from the literature of the direct costs of correctional education and of reincarceration. Focusing only on the direct costs of correctional education programs and of three-year reincarceration rates and using a hypothetical pool of 100 inmates, we estimated that the three-year reincarceration costs for those who did not receive correctional education would be between $2.94 million and $3.25 million. In comparison, for those who did receive correctional education, the three-year reincarceration costs are between $2.07 million and $2.28 million. This means that reincarceration costs are $0.87 million to $0.97 million less for those who receive correctional education. Given that the costs of providing education to this group of 100 inmates would range from $140,000 to $174,400, providing correctional education appears to be cost-effective when compared with the cost of reincarceration.

Another way to look at the cost-effectiveness of providing correctional education is to calculate the break-even point—defined as the risk difference in the reincarceration rate required for the cost of correctional education to be equal to the cost of incarceration. For a correctional education program to be cost-effective, we estimated that a program would need to reduce the three-year reincarceration rate by between 1.9 percentage points and 2.6 percentage points to break even. In fact, our meta-analytic findings indicate that participation in correctional education programs is associated with a 13 percentage-point reduction in the risk of reincarceration three years following release. Thus, correctional education programs appear to far exceed the break-even point in reducing the risk of reincarceration. Given that some programs appear more effective than others, the exact ratio of costs to benefits will naturally depend on the effectiveness of a particular program. Future investments in correctional education would ideally be designed in ways that allow for rigorous identification of effective programs' features.

CHAPTER FOUR

The Relationship Between Correctional Education and Employment

Introduction

This chapter presents the results from our meta-analysis where employment is the outcome. We first describe how we defined and measured employment across the 18 eligible studies, and we then pool all the studies together to provide an aggregate estimate of the relationship between participation in correctional education and employment. Next, we explore whether the relationship between correctional education and employment differs by the type of program and the method used to measure employment.

Measuring Employment

Employment was measured a number of ways across the 18 eligible studies along three dimensions: the *definition of employment* used by the researcher, the *time period* between release from prison and when employment is recorded for study participants, and the *statistical metric* used by the researcher to report differences in employment between the treatment and comparison group members. We describe each of these dimensions below in turn.

- *Definition of employment*. Employment is defined a number of ways, including having ever worked part-time since release, having ever worked full-time since release, employed for a specified number of weeks since release, and employment status (i.e., employed or not employed) at a particular time point. In our pool of 18 eligible studies, the most common way employment was operationalized was through a variable indicating whether the former inmate had ever worked full- or part-time since release (n = 9).
- *Time period*. Studies differed in the time period through which they followed the study participants after release from prison. Studies ranged from examining a cohort of former inmates in the community for three months since release from prison to following them for 20 years since release from prison. The most frequently used time period in the 18 eligible studies was one year (n = 7).
- *Statistical metric*. Fifteen of the studies simply reported the percentage or a weighted mean of the treatment and comparison groups that were employed, and three of the studies reported regression coefficients along with standard errors to express the magnitude of the difference in employment between the treatment and the comparison groups.

When there were multiple outcomes and reporting methods used, we gave preference to employment within one year of release or as close as possible to one year (as this represents the modal time period used by the authors of the studies) and regression coefficients (as this represents the best attempt by the authors of the studies to reduce potential sources of bias). However, as with our approach in our analysis of recidivism, we used whatever definition, time period, or statistical metric reported by the author so that we could be as inclusive as possible. As such, our employment measure comprises of a wide range of slightly different measures and thus should not be interpreted as any of the individual measures that make it up. Details on how each of the 18 studies defined and operationalized employment, as well as specific information on the individual programs being studied, the research design used in the study, the WWC Scale and the Maryland SMS ratings of the study's research design, and the rates of employment recorded for the treatment and comparison group, are shown in Appendix G.

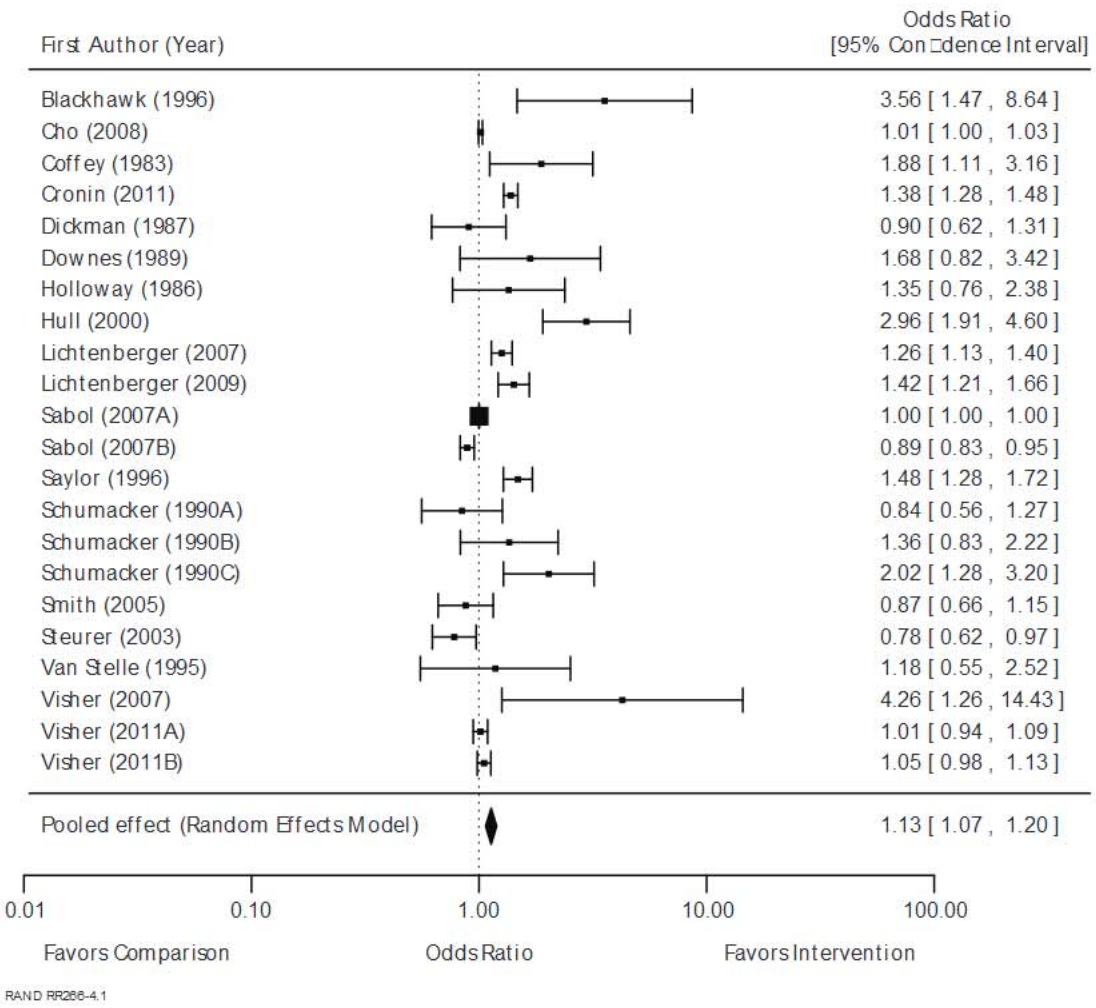
We transformed all 22 effect size estimates from the 18 studies into 22 odds ratios. Recall that the number of effect sizes exceeds the number of studies because a study could contain multiple treatment and comparison groups, and thus multiple comparisons. For our purposes, the odds ratio is calculated as the odds of obtaining employment among treatment group members divided by the odds of obtaining employment among comparison group members. Odds ratios greater than 1 indicate that the treatment group had a higher rate of employment, and odds ratios less than 1 indicate that the comparison group had a higher rate of employment. An odds ratio of 1 indicates that there is no difference between the treatment group and the comparison group.[1] These 22 odds ratios form the data points on which the random-effects regression is estimated.

Results: Estimates of the Relationship Between Correctional Education and Employment

To assess the relationship between correctional education and employment, we graphed the odds ratios for each of the 22 effect size estimates in Figure 4.1 using a forest plot. Similar to our analysis of recidivism, each row in the plot corresponds to an effect size, labeled on the left with the corresponding first author of the study and the year of publication. Studies with multiple effect size estimates are listed multiple times with a capital letter to differentiate among them. The black box represents the effect size for the study, and the "whiskers" extend to the range of 95 percent confidence intervals. The size of the box is proportional to the weight that is assigned to that effect size. The box and whiskers for each effect size are plotted in relation to the dashed line down the center of the graph, which indicates an odds ratio of 1. Effect sizes to the right of this line indicate that the treatment group had a higher odds of obtaining employment, and effect sizes to the left of this line indicate that the comparison group had a

[1] For example, in Lichtenberger's (2007) study of vocational education programs in Virginia correctional facilities, he determined that 71.5 percent of the treatment group found employment within 6.75 years of release and that 66.6 percent of the comparison group found employment within 6.75 years of release. The odds associated with a percentage of 71.5 are 0.715 / (1 – 0.715) = 2.51; in other words, the odds of a treatment group member obtaining employment are 2.51 to 1. The odds for the comparison group are 0.666 / (1 – 0.666) = 1.99; in other words, the odds of a comparison group member obtaining employment are 1.99 to 1. The associated odds ratio for this effect size is 1.26 (2.51 ÷ 1.99 = 1.26) and indicates that the odds of obtaining employment among treatment group members is 26 percent higher than the odds of obtaining employment among comparison group members.

Figure 4.1
Odds Ratios for Each of the 22 Effect Size Estimates



higher odds of obtaining employment. If the whiskers for the corresponding box do *not* cross this dashed line, then the study detected a significant difference between the treatment and comparison group for that particular effect size at the conventional level of $p < 0.05$.

The patterning of boxes and whiskers in Figure 4.1 shows that most studies report that the odds of obtaining employment are higher among the treatment group than the comparison group, as evidenced by most of the boxes corresponding to each size falling to the right of the dashed line. A small number of studies find a higher odds of obtaining employment in the comparison group, with two finding significant differences (Sabol, 2007; Steurer, Smith, and Tracy, 2003). The very last row displays the overall odds ratio for all 18 studies with 22 effect size estimates pooled together. The position of this overall odds ratio is indicated across the rest of the studies by the diamond at the bottom of the graph. The overall odds ratio is 1.13 ($p < 0.05$, 95 percent confidence interval = 1.07 to 1.20), indicating that across 32 years of empirical studies on the effects of correctional education, on average, the odds of obtaining employment postrelease among inmates receiving correctional education are 13 percent higher

than the odds of obtaining employment postrelease among inmates not receiving correctional education.

As with our analysis of recidivism, it is possible that the findings for employment favorable to correctional education programs may be driven by selection bias, wherein motivated, work-oriented inmates are selected (either by their own choice or by correctional program administrators) to enroll in educational programs. Therefore, the observed differences in employment between the treatment and comparison groups may reflect underlying differences in the types of inmates that participate in correctional education and not the causal effect of the program itself. To provide a better estimate of the potential causal relationship between program participation and employment, we recalculated the odds ratio for studies that fall at different levels of the Maryland SMS scale. The odds ratios and their corresponding confidence intervals are shown in Table 4.1. Ideally we would restrict our analyses to studies receiving a Level 4 or Level 5 rating on the Maryland SMS (as was done in our analysis of recidivism). However, as shown in this table, no studies with employment outcomes received a Level 5 rating and only one study received a Level 4 rating.[2] Therefore, we cannot test whether the positive relationship between correctional education participation and employment holds among studies with the most scientifically defensible research designs. Although we do detect an employment advantage favoring inmates receiving education while incarcerated, we cannot rule out selection bias as a potential explanation for this observed effect.

Interpreting the Relationship Between Correctional Education and Employment

As with our analysis of recidivism, we apply two other metrics to aid in interpretation: the *risk difference* and the *number needed to treat*. The *risk difference* is the absolute improvement in employment rates between those who received correctional education and those who did not. The *number needed to treat* indicates the predicted number of inmates who need to receive correctional education to secure one additional inmate postrelease employment. These two metrics require an estimated rate of employment in the population upon which to calibrate their calculations. Unfortunately, there is no national estimate of postrelease employment for former inmates that can serve this purpose. In lieu of a national estimate, we use the percent-

Table 4.1
Estimates of the Effect of Correctional Education Participation on the Odds of Postrelease Employment, by Levels of the Maryland Scientific Methods Scale

Maryland Scientific Methods Scale	Odds Ratio	95% Confidence Interval	n	k
Level 5	na	na	na	na
Levels 4 and 5	1.48*	1.28 to 1.72	1	1
Levels 3, 4, and 5	1.04	0.99 to 1.09	10	12
Levels 2, 3, 4, and 5 (total sample)	1.13*	1.07 to 1.20	18	22

*$p < 0.05$.

NOTE: n is the number of studies, k is the number of effect size estimates, and na is not applicable.

[2] The only study with employment outcomes receiving a Level 4 rating on the Maryland SMS is Saylor and Gaes' (1996) evaluation of the Post-Release Employment Project, which includes industrial work, vocational instruction, and/or apprenticeship training in federal prisons. They found that the treatment group yielded higher rates of employment after release (71.7 percent) than the comparison group (63.1 percent).

age of male inmates supporting themselves via employment at 15 months postrelease, based on a study of approximately 1,700 adult male inmates conducted between 2004 and 2007 in 12 states (Lattimore et al., 2012). We base our calculations on our odds ratio for those studies meeting a Level 3, Level 4, or Level 5 rating on the Maryland SMS, as these represent the highest-quality studies available to us. In this aforementioned multistate study, 66.0 percent of adult male inmates were employed at 15 months of release. Applying our pooled odds ratio, we find that *correctional education would be expected to improve postrelease employment rates by 0.9 percentage points.* Using these estimates, the number needed to treat (NNT) indicates that 114 inmates would need to receive correctional education to procure postrelease employment for one additional inmate.

Role of Program Type and Method Used to Collect Employment Data

We conclude our analysis of employment by exploring whether the relationship we observe between correctional education and the odds of obtaining employment varies varies by program type and/or the method used to collect employment data. The scientific review team abstracted both of these variables during their assessment and coding of the studies, which followed the review protocol shown in Appendix D. We use this information to recalculate our odds ratios separately for vocational programs and nonvocational programs and separately for studies that relied on administrative data, surveys to parole officers, and surveys to inmates. We focus on these two dimensions, because they have substantive and methodological implications for interpreting our main findings as well as for planning for future research in the field. Additionally, the data on these two variables are complete for our full sample of studies. Ideally, we would examine a broader range of program characteristics, but the data collected across studies were too inconsistent or incomplete. With a small pool of studies to examine, we consider these analyses to be purely exploratory. We urge readers to interpret these findings with that caveat in mind.

Program Type

In theory, vocational education programs should be more adept than traditional academic education programs at imparting labor market skills, awarding industry-recognized credentials, and connecting inmates with prospective employers. Therefore, we examine whether the relationship between correctional education and employment is stronger for vocationally oriented programs than traditional academic programs. To explore whether this is the case, we calculate odds ratios for effect size estimates corresponding to vocational programs and academic programs (combining ABE, high school/GED, postsecondary education programs) separately.[3] These odds ratios are presented in Table 4.2. Note that the summation of the number of studies in this table exceeds 18, because three studies contribute effect size comparisons for both vocational and academic comparisons. Although we might expect the relationship to be stronger for vocational programs, we find that both odds ratios for program type are greater than 1 and are statistically significant at $p < 0.05$. The odds ratio is higher for vocational programs than for academic programs, but they are not significantly different from one another—suggesting

[3] In our analysis of recidivism outcomes, we calculated odds ratios for ABE, high school/GED, and postsecondary education programs separately. Because of small sample sizes and our substantive focus on vocational programs, we combined these three programs into a single measure of "academic programs" for ease of interpretation and comparison.

Table 4.2
Estimates of the Effect of Correctional Education Participation on the Odds of Obtaining Employment, by Program Type

Program Type	Odds Ratio	95% Confidence Interval	n	k
Vocational education	1.28*	1.08 to 1.52	9	9
Academic education	1.08*	1.01 to 1.15	12	13

*p < 0.05.

NOTE: n is the number of studies and k is the number of effect size estimates.

that both academic and vocationally focused programs may be equally effective at preparing inmates for the labor market following release.[4]

Method Used to Collect Employment Data

Last, we explored whether the relationship between correctional education participation differed depending on the method used by the researcher to collect employment data. Most studies used state administrative data sources (n = 11), which measured only formal employment (i.e., jobs that are "on-the-books," such that the worker receives wages subject to tax withholding) within the state. Therefore, if the former inmates were self-employed, employed "under-the-table," or working in a state other than the one in which they were incarcerated, they were classified as not employed. Given that individuals with a criminal record are typically viewed less favorably by prospective employers and instead rely on nontraditional avenues for securing employment (Pager, 2003), it is possible the reliance on administrative records may understate employment gains made by correctional education participants. These limitations were overcome in studies that relied on surveys to parole officers (n = 5) or surveys to former inmates (n = 2) that inquired about postrelease employment histories. However, unlike state administrative data sources (which are typically complete), surveys are often hampered by low response rates and/or nonrandom response rates. The odds ratios for studies employing these different data collection methods are shown in Table 4.3.

Table 4.3
Estimates of the Effect of Correctional Education Participation on the Odds of Obtaining Employment, by Method Used to Collect Employment Data

Data Collection Method	Odds Ratio	95% Confidence Interval	n	k
Administrative records	1.07*	1.01 to 1.13	11	12
Survey to parole officer	1.61*	1.18 to 2.19	5	7
Survey to former inmate	1.04	0.94 to 1.16	2	3

*p < 0.05.

NOTE: n is the number of studies and k is the number of effect size estimates.

[4] A meta-regression shows that the ratio of the vocational odds ratio to the academic odds ratio is 1.09 (95 percent confidence intervals 0.98, 1.23; p = 0.125). Note that a meta-regression does not yield a direct ratio of the two corresponding odds ratios, which in the present case would be 1.19.

Studies that use administrative records and surveys to parole officers both find differences between treatment and comparison group members that are statistically significant at $p < 0.05$. However, the relationship between correctional education and employment is stronger in studies that use parole officer surveys than in studies that rely on administrative records: The odds ratio for parole officer surveys is larger than the odds ratio for administrative records (1.61 compared with 1.07), and their respective confidence intervals do not overlap. This suggests that in measuring only formal "on-the-books" employment, administrative records may potentially underestimate the effect of correctional education on labor force outcomes.

Summary

When examining 22 effect size estimates from 18 studies of correctional education programs spanning 32 years of research, the majority of studies we identified showed higher rates of employment among inmates receiving correctional education than among inmates who did not receive correctional education. On average, the odds of obtaining employment postrelease among inmates receiving correctional education are 13 percent higher than the odds of obtaining employment postrelease among inmates not receiving correctional education. No studies received a Level 5 rating and only one study receives a Level 4 rating. Therefore, we cannot rule out selection bias as a potential explanation for this observed relationship. Despite this limitation, our findings align with those produced in the meta-analysis by Wilson and colleagues (2000), which also found improved odds of employment among correctional education participants.

CHAPTER FIVE

The Relationship Between Computer-Assisted Instruction and Academic Performance

Introduction

This chapter presents the results from a meta-analysis in which standardized test scores in mathematics or reading are the outcome variables of interest, and in which the treatment variable of interest is correctional education administered via computer-assisted instruction rather than traditional, face-to-face classroom instruction. As noted in Chapter Two, only four studies that use achievement test scores met our eligibility criteria for inclusion. However, a benefit is that all four of these studies examine programs that use computer-assisted instruction—thus, allowing us to examine more closely an instructional delivery method that is increasingly popular in correctional settings. We first provide a brief description of the computer-assisted interventions themselves. As these studies are of clearly defined educational interventions (in contrast to most of the studies used in the recidivism and employment analyses), we describe them in detail to provide context for the results. We then describe how we standardized test scores across the four eligible studies. Next, we pool effect size estimates from the four studies to provide aggregate estimates of the relationship between computer-assisted instruction and students' academic performance in reading and mathematics. We then examine descriptive differences by program features. We conclude the chapter with a brief summary of key findings.

Description of the Computer-Assisted Instructional Interventions

All four of the studies discussed in this chapter compared computer-assisted instructional interventions to traditional, face-to-face classroom instruction led by a teacher. In each of the studies, the computer-assisted instruction replaced the same amount of time of traditional classroom instruction. All four studies were conducted in adult correctional education settings. In two of the studies—Batchelder and Rachal (2000) and McKane and Greene (1996)—students in both the treatment and comparison groups received additional, traditional classroom instruction beyond the portion of their instructional time that was subject to the intervention.

Two of the studies—Diem and Fairweather (1980) and Meyer, Ory, and Hinckley (1983)—assessed the same intervention—namely, the PLATO instructional software package for mathematics, reading, and language, published by PLATO Learning. This software was described as consisting of drill-and-practice instruction in basic skills that included arithmetic, reading, and language usage. In both studies, PLATO replaced face-to-face instruction led by a classroom teacher and covering similar content areas; in the Diem and Fairweather (1980)

study, the traditional classroom instruction was said to include "lecture, rote recitation, and some team teaching" (p. 207). The software was described as mastery-based and was supplemented by nonelectronic materials. The PLATO classrooms were staffed by a teacher and an aide in the Meyer et al. (1983) study and by a classroom teacher in the Diem and Fairweather (1980) study. In the Meyer et al. (1983) study, the intervention lasted approximately 2.5 hours per day for three months, at an implied rate of five days per week. In the Diem and Fairweather study, the intervention lasted eight weeks, but intensity and frequency were not specified.

The study by Batchelder and Rachal (2000) used a "tutorial/drill and practice" (p. 125) software package called Advanced Instructional Management System (AIMS) that allowed students to choose their focal areas and to progress at their own pace. It also provided diagnostic feedback on their progress. The software reportedly emphasized arithmetic and writing conventions, presenting students with lessons, sample problems to solve or essays to correct, feedback on their work, and chances to demonstrate learning from their mistakes. It was used to supplant face-to-face instructional time in mathematics, English, history, and science for one hour per day, five days a week, during a four-week period. AIMS classrooms were staffed by a facility employee rather than by a classroom teacher, and inmate peers were on hand to assist with technical difficulties.

McKane and Greene (1996) assessed the AUTOSKILL Component Reading Subskills Program (Fiedorowicz and Trites, 1987), which was reportedly designed to teach cognitive subskills of reading, and particularly syllable and word recognition. It offered speeded drill and practice and supplanted an unspecified portion of the traditional, teacher-led literacy instruction that the students otherwise received. Both AUTOSKILL and traditional instruction classrooms were staffed by literacy instructors. Traditional instruction was reported to include a variety of literacy teaching methods, including the Laubach method, Steck-Vaughn tutoring, peer tutoring, and traditional classroom instruction.

Notably, three of the four studies used random-assignment designs. Consequently, Bachelder and Rachal (2000) and Diem and Fairweather (1980) earned 5s on the Maryland SMS, and McKane and Greene (1996) earned a 4 due to high attrition. The other study, Meyer, Ory, and Hinckley (1983), did not take steps to reduce selection bias and thus earned a 2 on the Maryland SMS.

Measuring Academic Performance

For the meta-analysis, we limited our examination of academic performance to the two content areas that were common to more than two studies—namely, mathematics and reading. These are policy-relevant measures, since they are building-block skills for other content areas, and they are the two subjects that states are required to measure annually in public schools under the federal No Child Left Behind Act of 2001. Beyond these content areas, one study also included a language test (Meyer et al., 1983) and another included measures of vocabulary and spelling (Diem and Fairweather, 1980), but to include an outcome variable in the meta-analysis, we required at least three studies to measure that variable.

Each study employed one of three commercially available standardized tests to measure academic performance. All were paper-and-pencil examinations, and all used separate pretests and posttests to measure changes in student performance over time. Information provided about the standardized tests is described below.

One study (Diem and Fairweather, 1980) used the *Adult Basic Learning Examination (ABLE), Level II*, which is designed to measure the performance of adult students performing on a fifth- to eighth-grade level. Our analysis focused on the subscale scores in reading and total arithmetic; the latter comprises computation and problem-solving subscales.

One of the studies (Batchelder and Rachal, 2000) used the Comprehensive Adult Student Assessment System (CASAS) mathematics and reading scales. This test is reportedly designed to measure performance from beginning levels through high school completion and was reportedly "validated through field testing based on 15 years of assessment data from more than 2 million adult learners" (Batchelder and Rachal, 2000, citing the Comprehensive Adult Student Assessment System, 1996).

The other two studies used the Test of Adult Basic Education (TABE) scales in reading (Meyer, Ory, and Hinckley, 1983) or mathematics and reading (McKane and Greene, 1996). Meyer, Ory and Hinckley (1983) used the TABE M (medium level) as a pretest and TABE D (difficult level) as a posttest. The former is reportedly designed to reliably measure performance at grades 3 through 10 and the latter, at grades 5 through 12. McKane and Greene (1996) did not specify the versions used, but both studies noted that the TABE is frequently used as a measure of academic performance in correctional settings.

Creating a Common Performance Scale

To synthesize the results of studies that use different measures of academic performance with different testing scales, it is necessary to put the results in common units across studies. Many studies and research syntheses have to create a common scale across disparate tests by converting scores to standard deviation units or z-scores, where a standard deviation is defined as the average deviation from the mean across test-takers on a given assessment.[1] In this case, however, all of the test scores are reported in grade equivalents or in forms that can be easily converted to grade equivalents, so we use these as our common metric, thereby avoiding the need to use standard deviation units for different tests (Baguley, 2009).[2] Grade-level equivalents have the additional benefit of being easily understood by policymakers and practitioners, because one unit is equal to a single, nine-month academic year of learning in a particular content area. This metric typically refers to a standard scholastic setting rather than a correctional education setting, in which students receive approximately one hour of instruction in each of six to seven content areas for five days per week. As such, one month of learning (as reported on the ABLE, for instance, in Diem, 1980) would represent one-ninth of a grade-level equivalent. According to a publicly available report from the CASAS (2012), four scale score points on both the reading and mathematics scales represent a one-grade level difference. Consequently, we defined a unit difference in CASAS score points as equal to one-quarter of a grade-level equivalent. For the two studies that used the TABE, results were already presented in terms of grade-level equivalents. Because we were able to transform ABLE and CASAS scores linearly into grade-

[1] More technically, a standard deviation is the square root of the squared deviation from the mean, divided by n – 1.

[2] Moreover, only two of the four studies (Batchelder and Rachal, 2000; and Meyer, Ory, and Hinckley, 1983) reported standard deviations of student performance. The other two reported only standard deviations of student performance changes, and deviations for an appropriate comparison population were not publicly available for the ABLE, in particular.

level equivalents, and because TABE scores were already reported in grade-level equivalents, we were able to report effects consistently across studies using this metric.[3] Additional details about how each of the four studies defined and operationalized achievement, as well as specific information on the individual interventions, the research design used in the studies, the WWC and Maryland SMS ratings, and the test scores for the treatment and comparison groups, are shown in Appendix H.

Results: Effects of Computer-Assisted Correctional Education on Student Performance in Math and Reading

The four aforementioned studies include a total of nine effects. Three of the studies provide one math effect and one reading effect each, and one of the studies (McKane and Greene, 1996) contributes no math effect but does contribute separate reading effects for three distinct subgroups—students beginning at the third-grade reading level or lower, students beginning between the third- and sixth-grade levels, and students beginning above the sixth-grade level. In the studies that include both reading and mathematics estimates, there is complete overlap between the samples of reading and mathematics test-takers, meaning that the estimates for each content area are not independent within a given study. As a result, we present separate meta-analytic estimates for reading and mathematics rather than combining the estimates into a single academic achievement effect.

These effect estimates for reading are summarized in a forest plot shown in Figure 5.1, and the estimates for mathematics are also shown in Figure 5.2. In each plot, the horizontal axis represents the estimated effect of computer-assisted instruction relative to traditional instruction. As noted above, the effect estimates are denominated in grade-level equivalents, so that one unit corresponds to a single grade level of learning, or approximately the knowledge that would be gained in nine months of full-time classroom instruction, on average. For each study listed on the left of the figures, the black box represents the effect size estimate for a given study sample or subsample, and the size of the box is proportional to the size of the sample or subsample. The horizontal line for each study represents the 95 percent confidence interval around the effect.[4] Each individual effect and its confidence interval are also listed in the right-hand column of the figure. The overall, meta-analytic effect across studies is estimated as in prior chapters with a random effects regression analysis, which weights each effect according to its sample size and the precision with which it is estimated.

---

[3] The actual analysis uses scale-score units. In two of the studies, scale scores and standard deviations are provided for both the pretest and posttest scores. One of these studies (Batchelder and Rachal, 2000) provides an F-test on the posttest difference, from which we back out a standard error, so the meta-analysis includes only the posttest difference for that study. The other of these studies (Meyer, Ory, and Hinckley, 1983) provides p-value thresholds for the pre-post differences in each group; we back out the standard errors using the most conservative assumptions for these p-value thresholds. The other two studies (Diem and Fairweather, 1980; McKane and Greene, 1996) provide standard errors for the pre-post difference in scale scores of each group, and we use those standard errors in the analysis. In other words, the meta-analysis uses the pre-post differences in scale scores for each group (and associated standard errors) for all of the studies except Batchelder and Rachal (2000), where we instead include only the post-test difference and associated standard error.

[4] Note that the right whiskers for McKane (1996a) and Batchelder (2000b) are arrows. This is to signify that the confidence intervals for these effect sizes extend beyond the scales of the figures.

Figure 5.1
Reading Effect Estimates

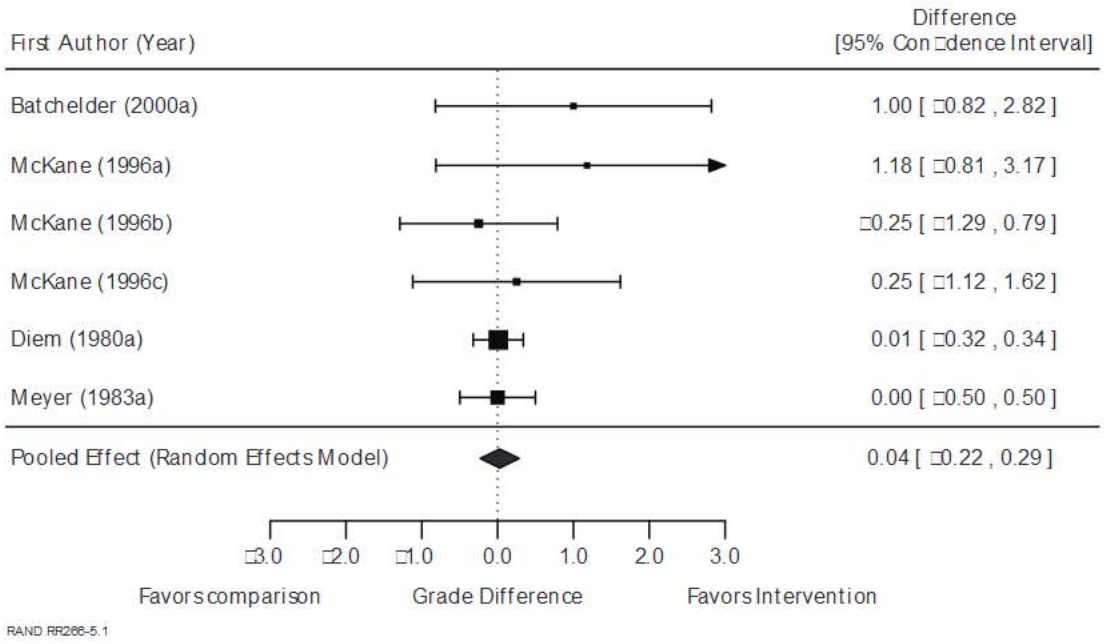
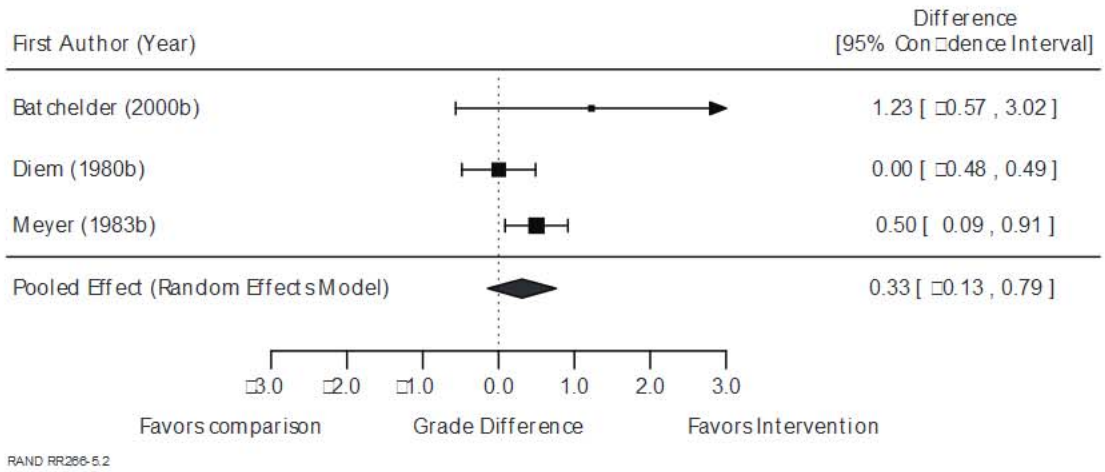


Figure 5.2
Mathematics Effect Estimates



As shown in the bottom row of Figure 5.1, we estimate that the overall effect of computer-assisted instruction relative to traditional instruction in reading is 0.04 grade levels, or about 0.36 months of learning. This is a small effect in substantive terms and is also not statistically distinguishable from zero, as evidenced by the 95 percent confidence interval, which ranges from –0.22 to 0.29. The fact that zero falls within the confidence interval means that we cannot reject the null hypothesis that computer-assisted instruction offers no benefit in reading beyond that of traditional instruction.

Turning to Figure 5.2, we estimate a substantively larger effect of computer-assisted instruction on achievement in mathematics. There, we find an effect estimate of 0.33 grade levels, which represents about three months of learning. Taken at face value, this is a substantial effect, particularly given that the dosages ranged from only one month of instruction (at one hour per day) in the case of Batchelder and Rachal (2000) to two months in the case of Diem and Fairweather (1980) and three months (2.5 hours per day) in the case of Meyer, Ory, and Hinckley (1983). Assuming that a standard deviation in the outcome is about 1.5 grade-level equivalents (based on estimates from Meyer, Ory, and Hinckley, 1983), this represents about a fifth of a standard deviation. To put the finding in context, this effect size estimate is roughly twice what many studies find to be the difference in effect between a high-performing and an underperforming teacher (e.g., Aaronson et al., 2007; Kane and Staiger, 2005; Rivkin et al., 2005; Rockoff, 2004).[5] The estimate is based on only three studies. In light of the limited number of studies and the limited number of participants within each, the 95 percent confidence interval around the estimate ranges from –0.13 to 0.79 grade levels. As is true for reading, the fact that zero falls within the confidence interval means that the result is not statistically significant at the 5 percent level. We therefore fail to reject the null hypothesis that computer-assisted and traditional instruction have identical effects on student performance in mathematics.

Viewed from another perspective, however, the data also provide no evidence that computer-assisted instruction harms student performance. Because computer-assisted instruction can be self-paced and can be supervised by a person other than a licensed classroom teacher, it is potentially less costly to administer and could even allow correctional facilities to expand their instructional course offerings. For these reasons, the finding of no statistically significant difference between computer-assisted and face-to-face instruction suggests that, based on current evidence, computer-assisted instruction may be a reasonable alternative to traditional, face-to-face classroom instruction in correctional facilities. Moreover, the most recent of the four studies in our meta-analysis that addressed this question was published in 2000, and two were published in the early 1980s. The capability and utility of computer-assisted instructional technology has progressed substantially since these studies were published (U.S. Department of Education, 2010). It is possible that the effects of newer technologies could outstrip those found in the studies described here. Therefore, it will be important for such technologies to be carefully evaluated when they are deployed in correctional settings.

Role of Program Type

Practitioners may also wonder about the extent to which one type of computer-assisted instruction outperforms another. To address that question, we conclude our analysis of achievement by exploring whether the relationship we observe between computer-assisted instruction and learning in correctional facilities varies by program type. Table 5.1 presents details about the program type associated with each effect estimate.

Again, we lack enough studies to address this question formally, but examining Table 5.1 does yield some descriptive information about differences by intervention type. Two of the studies—Diem and Fairweather (1980) and Meyer, Ory, and Hinckley (1983)—used the PLATO drill-and-practice software relative to regular classroom instruction, whereas

[5] Where the difference is one standard deviation of teacher effectiveness.

Table 5.1
Estimates of the Effect of Computer-Assisted Instruction on Student's Achievement Grade Level, by Content Area and Program Type

Study	Content Area	Program Type	Effect Estimate	95% Confidence Interval
Batchelder (2000a)	Reading	AIMS	1	−0.82 to 2.82
McKane (1996a)	Reading—low baseline	AUTOSKILL	1.18	−0.81 to 3.17
McKane (1996b)	Reading—medium baseline	AUTOSKILL	−0.25	−1.29 to 0.79
McKane (1996c)	Reading—high baseline	AUTOSKILL	0.25	−1.12 to 1.62
Diem (1980a)	Reading	PLATO	0.01	−0.32 to 0.34
Meyer (1983a)	Reading	PLATO	0	−0.50 to 0.50
Batchelder (2000b)	Mathematics	AIMS	1.23	−0.57 to 3.02
Diem (1980b)	Mathematics	PLATO	0	(−0.48 to 0.49)
Meyer (1983b)	Mathematics	PLATO	1.04*	(0.09 to 0.91)

*$p < 0.05$.

NOTES: The Study column lists only the first author and year for each study. The full citation for each study can be found in Appendix H.

Batchelder and Rachal (2000) used a software package called AIMS that focused on basic arithmetic skills and writing conventions, and McKane and Greene (1996) used the AUTOSKILL syllable-and-word-recognition software. Turning first to the results from the two PLATO studies, we find that they are uniformly close to zero except for the mathematics effect in Meyer, Ory, and Hinckley (1983), where there is a significant and positive effect of half a grade level, or about 4.5 months. This is substantial, since the intervention lasted only three months.[6] The largest effects we see are from the Batchelder and Rachal (2000) study, where we find 20 hours' worth of computer-assisted instruction with AIMS arithmetic and language practice software yielding effects of more than a single grade level in both math and reading. However, these effects have very large confidence intervals, rendering them statistically nonsignificant, and, unlike results from the other three studies, they are unadjusted for substantial baseline differences at pretesting because the correlation between pretest and posttest scores was not reported.

Finally, we turn to McKane and Greene (1996), whose results seemed to depend on the baseline reading ability of students. For students who began with lower than a third-grade reading level (effect "a") the syllable-and-word-recognition software was associated with gains of more than one full grade level, although the sample was small and the result was not statistically significant.[7] For students with baseline reading levels between grades three and six (effect "b") or above grade six (effect "c"), the results were either negative or slightly positive and were nonsignificant in all cases. In sum, the data are slightly positive with regard to PLATO effects in mathematics and AIMS in both math and reading, and AUTOSKILL only for the lowest-

[6] The intensity of 2.5 hours per day of instruction is comparable to what a student might receive in math and language arts alone in a traditional secondary school environment, which is the environment on which grade-level equivalents are based.

[7] Note that the duration and frequency of the intervention were not reported.

skilled individuals. However, with so few studies of each intervention, our ability to generalize about any given intervention is quite limited.

Summary

Our meta-analyses of six reading effect estimates and three mathematics effect estimates from four studies suggest that the effect of computer-assisted instruction on incarcerated adults' reading and mathematics performance is not statistically different from that of traditional, face-to-face classroom instruction. The overall effect of computer-assisted instruction is estimated at only about 0.36 months of learning in reading but at a more substantial three months of learning in mathematics. Although the mathematics effect estimate is substantively meaningful, its confidence interval includes zero and, thus, we cannot rule out the possibility that it is due to chance alone. Moreover, as none of the prior meta-analyses on correctional education looked specifically at computer-assisted instruction and achievement, our findings cannot be directly compared with existing work in this area.

CHAPTER SIX

Conclusions

The goal of this report was to address the question of what we know about the effectiveness of correctional education—academic programs and vocational training programs—for incarcerated adults in U.S. state prisons. Specifically, we examined the evidence about the relationship between correctional education and recidivism and postrelease employment outcomes and the relationship between academic performance and computer-assisted instruction. These findings will inform policymakers, educators, and correctional education administrators interested in understanding the association between correctional education and reductions in recidivism and improvements in employment and other outcomes.

In this chapter, we summarize our overall findings, provide specific recommendations for strengthening the evidence base in this field, and discuss the policy implications and next steps.

Overall Summary of Findings

Our meta-analytic findings provide additional support to the premise that receiving correctional education while incarcerated reduces an individual's risk of recidivating after release. After examining the higher-quality studies,[1] we found that, on average, inmates who participated in correctional education programs had *43 percent lower odds of recidivating* than inmates who did not. These results were consistent even when we included the lower-quality studies in the analysis. This translates as a reduction in the risk of recidivating of 13 percentage points for those who participate in correctional education programs versus those who do not. This reduction in the risk of recidivating is somewhat greater than that reported by Wilson, Gallagher, and MacKenzie (2000), which showed an average reduction in recidivism of about 11 percentage points. Using more recent studies and ones of higher quality, our findings complement the results published by Wilson, Gallagher, and MacKenzie (2000), Aos, Miller, and Drake (2006), and MacKenzie (2006) and provides further support to the assertion that correctional education participants have lower rates of recidivism than nonparticipants.

Given the high percentage of state prison inmates who have not completed high school, participation in high school/GED programs was the most common approach to educating inmates in the studies we examined. We found that inmates who participated in high school/GED programs had a 30 percent lower odds of recidivating than those who had not. In gen-

[1] That is, RCTs or quasi-experimental designs where the treatment and control groups are matched at baseline on at least three characteristics other than gender.

eral, studies that included ABE, high school/GED, postsecondary, and/or vocational training programs showed a reduction in recidivism. However, it is not possible to disentangle the effects of these different types of educational programs, because of the overlap in curricular exposure and a lack of specificity about dosage. Thus, we cannot assert, for example, that high school/GED programs have a greater effect on reducing recidivism than postsecondary education programs.

When we look at the relationship between correctional education and postrelease employment, our meta-analyses found—using the full set of studies—that the odds of obtaining employment postrelease among inmates who participated in correctional education (either academic or vocational programs) was 13 percent higher than the odds for those who did not. However, only one study fell into the higher-quality category.[2] Thus, if one wants to base policy decisions on the higher-quality studies alone, then we are limited in our ability to detect a statistically significant difference between program participants and nonparticipants in postrelease employment. Still, our results suggest a positive association between correctional education and postrelease employment. This finding aligns with those produced in the Wilson, Gallagher, and MacKenzie (2000) meta-analysis, which also found improved odds of employment among correctional education participants.

When examining the relationship between correctional education and postrelease employment, one might expect vocational training programs to be more adept than academic education programs at imparting labor market skills, awarding industry-recognized credentials, or connecting individuals with prospective employers. And, indeed, when we looked at the relationship between vocational training—versus academic correctional education programs—and postrelease employment, we found that individuals who participated in vocational training programs had odds of obtaining postrelease employment that were 28 percent higher than individuals who had not participated in vocational training. In comparison, individuals who participated in academic programs (combining ABE, high school/GED, and postsecondary education programs) had only 8 percent higher odds of obtaining postrelease employment than individuals who had not participated in academic programs. Although the results suggest that vocational training programs have a greater effect than academic programs on one's odds of obtaining postrelease programs, there was no statistically significant difference between the odds ratios for the two types of programs.

We also examined the relationship between computer-assisted instruction and academic performance—something that was not examined in any of the previous meta-analyses. In this case, the outcomes of interest were standardized test scores in mathematics or reading. We reviewed four studies[3] that compared the achievement test scores of inmates receiving computer-assisted instruction with the achievement test scores of inmates receiving face-to-face instruction. In two of the studies, students in both the treatment and comparison groups also received additional, traditional classroom instruction beyond the portion of their instructional time that was computer-assisted. We limited our examination of academic performance to the two content areas that were common to more than two studies—math and reading.

---

[2] This study by Saylor and Gaes (1996) examined industrial work, vocational instruction, and apprenticeship in federal prisons and found a 71.7 percent higher rate of employment among those who participated in these programs compared with 63.1 percent for those who had not.

[3] Three of these four studies employed high-quality research designs as defined by the WWC rating scheme and the Maryland SMS.

We estimated that the overall effect of computer-assisted instruction relative to traditional instruction is 0.04 grade levels in reading, or about 0.36 months of learning, and 0.33 grade levels in mathematics, which represented about three months of learning. In other words, on average across the study samples, students exposed to computer-assisted instruction learned very slightly more in reading and substantially more in mathematics as compared to those exposed to traditional instruction for the same amount of instructional time. However, these differences were not statistically significant and thus may be due to chance alone.

Because computer-assisted instruction can be self-paced and can be supervised by a tutor or an instructor, it is potentially less costly to administer than traditional instruction. It is worth noting that since the publication of these four studies,[4] the capability and utility of instructional technology has progressed (U.S. Department of Education, 2010), which suggests that the effects of the newer technologies may potentially outstrip those found in the studies examined here. The current positive (though not statistically significant) result, the potential cost-effectiveness of computer-assisted technology, and the fact that the technology is getting better suggest that its use in this context could be promising.

State policymakers, corrections officials, and correctional education administrators are asking a key question: How cost-effective is correctional education? In other words, although our findings clearly show that providing correctional education programs is more effective than not providing them, such programs have costs. Thus, to place our meta-analytic findings into context, we undertook a cost analysis using estimates from the literature of the direct costs of correctional education programs and of incarceration itself, and using a three-year reincarceration rate. Our estimates show that the direct costs of providing education to a hypothetical pool of 100 inmates would range from $140,000 to $174,400 with three-year reincarceration costs being between $0.87 million to $0.97 million *less* for those who receive correctional education than for those who do not. This translates as a per inmate cost ranging from $1,400 to $1,744, suggesting that providing correctional education is cost-effective compared with the cost of reincarceration. We also calculated the *break-even point*—defined as the risk difference in the reincarceration rate required for the cost of correctional education to be equal to the cost of incarceration. For a correctional education program to be cost-effective, we estimated that a program would need to reduce the three-year reincarceration rate by between 1.9 percentage points and 2.6 percentage points to break even. In fact, our meta-analytic findings show that participation in correctional education programs is associated with a 13 percentage point reduction in the risk of reincarceration three years following release from prison. Thus, correctional education programs appear to far exceed the break-even point in reducing the risk of reincarceration.

Our analysis focused only on the direct costs of correctional education programs to the prison system. A full analysis of the benefits and costs of correctional education besides accounting for the direct costs to a prison system would also need to account for other costs, such as the financial and emotional costs to victims of crime and to the criminal justice system as a whole, which could be much more substantial than our estimates above. The Washington State Institute for Public Policy's (WSIPP) undertook a cost-benefit analysis for its state comparing different types of adult rehabilitative programs, including education programs. Using a conservative set of assumptions, WSIPP found that vocational training and general education

[4] Two of the studies were published in the early 1980s; the other two were published in 2000.

in prison produced some of the largest net economic benefits for adult programs (Aos, Miller, and Drake, 2006).

Last, in considering the above findings, it is important to keep in mind that the 2008 recession also had an effect on the field of correctional education. The recession affected correctional education (and other rehabilitative) programs in a number of states, leading to some dramatic changes in the number of programs offered, the sizes of classes, the modes of delivery, and the number of inmates who participate in these programs. For example, funding for correctional education was reduced by 30 percent as part of California's $1.2 billion budget reduction for corrections in fiscal year 2009 (California Rehabilitation Oversight Board, 2010). As a result, approximately 712 teaching positions were eliminated, the number of vocational programs was reduced by nearly 50 percent, and the capacity of academic and vocational programs was reduced by 3,300 and 4,500 slots, respectively. To reduce the effect of these cuts on capacity and to maximize enrollment, the California Department of Corrections and Rehabilitation also developed five new education models with decreased program frequency, duration, and options while maximizing the number of inmates with access to the programs. For example, under the new education models students would meet for three hours per day once a week (which would allow for two sessions during the day) instead of meeting for 6.5 hours a day, five times a week under the old education model.

In Texas, the legislature reduced the budget for its state prison education system by approximately 27 percent, or $17.8 million per year over the next biennium (Windham School District, 2011–2012). To address the reduction in funding, 271 full-time equivalents (FTEs) were eliminated, all staff received reductions in salary, and other cuts were implemented (e.g., to supplies, travel, and other operating budgets).

In Oklahoma, budget cuts affected both academic and vocational programs. For example, appropriations to CareerTech (which runs the state prison Skills Centers that provide vocational and technology training) declined by more than 15 percent between fiscal years 2009 and 2012 (Wertz, 2012). Five of the state's 15 prison Skills Centers were closed, resulting in the loss of vocational training capacity in welding, carpentry, masonry, plumbing, and electrical. Since 2008, the Oklahoma Department of Corrections lost one-third of its full-time education staff and a similar percentage of its Skills Center instructors.

Within the past year, there has been an uptick in funding for correctional education, with many state correctional education directors reporting either no further funding cuts or even some minor increases in funding—a situation that has enabled them to begin modestly rebuilding programs (personal communication, Correctional Education Association [CEA] Leadership Forum, 2012). That said, a reduced funding environment will likely be true for correctional education programs for the near future, and the return on investment of these programs will likely continue to be a topic in state-level budget discussions.

The Need to Improve the Research Evidence Base for Correctional Education

Using the most recent published studies in the field, we similarly find that the quality of the available research on correctional education is highly variable (Gaes, 2008; MacKenzie, 2008). Unlike authors of previous meta-analyses, we had more studies with which to assess the effectiveness of correctional education. However, although our meta-analyses, as did previous meta-analyses, accounted for the strength of the research designs of the various studies examined,

there are still a number of questions of interest to educators and policymakers that the current literature—with its variable research quality—does not permit us to address. For example, we would want to look "inside the black box" of correctional education programs to try to understand what program elements (e.g., types of curriculum, mode of instruction, dosage, type of instructors) are associated with effective programs with respect to reductions in recidivism and improvements in postrelease employment outcomes.

In addition, one would want to address such questions as:

1. What dosage is associated with effective programs and how does it vary for different types of students?
2. Who benefits most from different types of correctional education programs?
3. What types of correctional education programs are associated with the highest postrelease returns?
4. What factors moderate or mediate the effect of correctional education?
5. How effective are peer tutors compared with credentialed instructors?
6. What is the right balance between in-person instruction versus self-study or computer-based learning?
7. What principles from adult education and learning may be applicable to correctional education?

All these questions get at the need to improve the evidence base. Below we provide recommendations for improving the evidence base in four critical areas:

1. Apply stronger research designs.
2. Measure program dosage.
3. Identify program characteristics.
4. Examine more proximal indicators of program efficacy.

Applying Stronger Research Designs

As discussed in this report, establishing a causal relationship between correctional education participation and successful outcomes for inmates requires ruling out the possibility of selection bias. This form of bias occurs when inmates who elect to participate in educational programs differ in unmeasured ways from inmates who elect not to participate in educational programs. In other words, correctional education participants may be more motivated, have a stronger internal locus of control, be more proactive about planning for their postrelease futures, etc.—all of which could affect why participants do better, independent of the effect of the programs themselves. Thus, if such differences between the treatment and comparison group exist before participation, any observed postparticipation outcomes may not necessarily reflect the causal effect of the program. In other words, higher rates of employment and lower rates of recidivism among correctional education participants may reflect inmates' skills and temperament and have nothing (or little) to do with exposure to education while incarcerated. Isolating the effects that can be directly attributable to the program itself is crucial in supporting the design of effective policies—an objective that is hampered by studies with research designs that are highly susceptible to selection bias.

[illegible] studies used to assess recidivism and one [illegible] of the 50 studies [illegible] in our meta-analysis, only several met [illegible] based on studies that received a Level 5 rating (a [illegible] 18 studies used to assess employment were ba

well-executed RCT) or a Level 4 rating (a quasi-experimental design with very similar treatment and comparison groups) on the Maryland SMS. Most of the studies were based on lower-quality research designs (Level 3 and below on the Maryland SMS) that were susceptible to selection bias. Further, many studies did not report sufficient information about the sociodemographic characteristics and other characteristics of the treatment and comparison groups; reporting on such information would allow for meaningful differences between the two groups to be evaluated and the potential threat of selection bias to be quantified.

To minimize this potential for bias, future studies should ideally employ research designs that that help to minimize it. The ideal design, of course, is an RCT, in which individuals are randomly assigned to the treatment group (e.g., those who receive vocational training) and to the control group (those who do not); however, RCTs may not always be practical or politically feasible with a criminal justice population.

When an RCT is not possible, two other alternatives might be feasible—a regression discontinuity (RD) design and a propensity score matching/weighting design. Both alternatives are intended to minimize selection bias, although an RD design does so more rigorously, because it addresses selection on *both* unobserved and observed attributes, whereas propensity scores address only the latter. The RD design, when executed properly, would merit a Level 5 on the Maryland SMS, in keeping with WWC standards for RDs (Schochet et al., 2010), whereas a propensity score matching or weighting study would merit a Level 4 rating at best.

Using an RD approach, assigning inmates to the treatment group would be based on a strict cut-point from a continuous measure that is judiciously applied to every inmate. For example, scores on the TABE may be used to select inmates to participate in a correctional education program, such that everyone directly above the cut-point is assigned to the program (i.e., to the treatment group) and everyone below the cut-point is assigned not to receive the program (the control group).

A key assumption of the RD design is that there is a linear relationship between the selection mechanism and the outcome, or that the relationship can be linearized. If this assumption holds and the design is properly implemented, then this design has very high internal validity. Because the assignment rule is fully understood and modeled, assignment is removed from the estimate of the treatment effect. To be implemented well, an RD design requires reasonably strong compliance with the assignment rule, although effects can be scaled for partial noncompliance through an instrumental variable analysis. It is noteworthy that none of the studies in our meta-analyses used an RD design.

Propensity score matching or weighting is possible when there is a range of information collected on program participants and nonparticipants—including sociodemographic information, prior criminal records, prior education and labor force experiences, cognitive functioning, and, if possible, other personality and behavioral traits. This information can be used to create a comparison group that is evenly balanced with the treatment group on the observed set of characteristics maintained in the data. In doing so, those in the comparison group have approximately the same "propensity" to have enrolled in correctional education as those in the treated group. This matching or weighting helps attenuate the threat of selection bias when making comparisons on the outcomes of interest, particularly when the set of characteristics used to balance the treatment and comparison groups is extensive and includes variables most likely to differentiate participants from nonparticipants.

Among the studies in our meta-analyses, only five used propensity score matching or weighting, although many more (virtually all the Level 4 and some of the Level 3 studies) used

a manual matching procedure in which treated inmates were matched to similar untreated inmates on key variables using administrative records. As long as the studies showed baseline equivalence between treatment and comparison groups on age, prior offenses, baseline educational level, and time between release and data collection, they were assigned a Level 4, regardless of the matching procedure used. However, the advantage of propensity score matching over a manual matching procedure is that it can help researchers obtain baseline equivalence over a much larger number of variables than one can typically achieve with a manual matching procedure (Rosenbaum and Rubin, 1983).

In addition, identifying the appropriate comparison groups is important (Gaes, 2008). Many of the studies reviewed in our meta-analyses used comparison groups of nonprogram participants but did not consider differences in terms of levels of education, certification, or training. As a result, the comparison group might be a mixture of inmates with varying levels of academic achievement.

Gaes (2008) recommended that a study registry be established to help sort out the different effect sizes found across studies. Given the vast array of programs currently administered, the dearth of basic information on their design and their effectiveness in a centralized system precludes the effective utilization of resources—particularly for states making strategic decisions on whether and how to recalibrate their programs to adjust to changes in funding and changes in the prisoner population. Such a registry would include details about each study including information about the program and intervention, about the evaluation design, characteristics of the treatment and comparison groups, and outcomes measures used. The research summaries provided in Appendixes F, G, and H of this report serve as a potential template for this type of information. The methodological and dissemination approach used by the U.S. Department of Education's What Works Clearinghouse could be adopted and adapted by the corrections' community to serve as a registry whereby programs are systematically cataloged and reviewed—improving accountability on the part of the programs and building a high-quality research base that can help better inform questions regarding what works best to effectively educate incarcerated individuals.

Measuring Program Dosage

Many practitioners have posed the question, What dosage level is associated with effective correctional education programs? For instance, does it matter that an individual participates in 20 hours of academic instruction or is 30 hours of academic instruction required for a given course? In other words, how much correctional education is needed to be effective? Such questions of dosage are especially salient now, when many correctional education programs have experienced significant budget cuts.

On average, the studies we reviewed lacked specific information about the dosage of the program, such as the overall program duration, the number and grade level of the courses in which inmates were enrolled, how many hours per day or week inmates were exposed to formal class instruction, and how many hours per day or week inmates worked on assignments outside the classroom. In many of the studies, particularly those that were secondary analyses of administrative data sets, respondents were categorized simply as correctional education participants and nonparticipants. This crude categorization undoubtedly masked variation in exposure to the program among participants. For example, some inmates may have been enrolled for a year; others may have been enrolled for a week and withdrawn.

Without being able to discern such differences, it is difficult to put the findings from individual studies in their proper contexts. Some studies may have produced null findings, not because the program was ineffective if implemented as designed but because the average dosage that the treatment group received was too small to make a difference. The lack of dosage information means that there is little to no empirical evidence that can help inform policymakers on how much correctional education is necessary to produce a change in the desired outcomes. In future studies, the proper recording of program dosage when collecting data and monitoring the progress of inmates through correctional programs will be critical to enable researchers to examine program dosage.

Identifying Program Characteristics

When we undertook our review of the literature, our charge from BJA was to identify promising or evidence-based programs that could be potentially replicated in other settings. We were limited in our ability to do so, because many of the studies did not provide sufficient detail on the characteristics of the program, such as the structure of the curriculum, the training and certifications of the teachers, the instructional methods used by the teachers, the student-teacher ratio in classrooms, and supplemental access to textbooks and technology.

To the extent possible, we culled this information from the studies that provided it and used it in an exploratory fashion in our meta-analyses. However, few studies consistently listed these details in their program descriptions; consequently, our findings from these few studies are suggestive at best. Thus, from a meta-analytic approach, we are unable to offer evidence-based prescriptions about what aspects of correctional education are most or least effective. The field would be well served if future research carefully documented the characteristics of the programs so that different models of program organization and instruction could be empirically validated.

Examining More-Proximal Indicators of Program Efficacy

The majority of studies used recidivism as an outcome measure. However, some would argue that recidivism is a *distal* measure that can be affected by many factors beyond correctional education. Further, studies differ in how recidivism is measured and in the length of time that recidivism is tracked.

Instead, many would argue that what is needed are more proximal measures that would better indicate how programs actually affect thinking and behavior, such as changes in motivation, literacy gains, development of concrete skills, or academic progress versus academic achievement.

The overwhelming number of studies we reviewed used recidivism as the major indicator, which is understandable given its importance as a marker of successful prisoner rehabilitation. However, despite its salience in criminological research, the emphasis on recidivism leaves much less known about the process through which correctional education helps shape how former inmates reintegrate into the community. Correctional education is believed to improve the skills and abilities of inmates (i.e., "human capital" in economics parlance), which, in turn, improves their chances of continuing education/training upon release and then finding gainful employment.

Only four studies in our review looked at skills and abilities (as measured by achievement test scores) and only 18 looked at employment. There were too few studies of additional education/training to include in a meta-analysis. Applying these more proximal indicators of

program efficacy will help to better elucidate the mechanisms that undergird the role of education in the rehabilitation process.

For example, collecting information on cognitive gains while inmates are enrolled in the program, additional education and training that inmates receive following their release, and more-detailed information about their postrelease employment (e.g., timing of employment, method of hiring, wages, occupation type, sector) would be important. Additionally, with respect to employment, our analysis and other research studies recognize that solely relying on administrative records, which record only formal "on-the-books" jobs, may underestimate the effect of correctional education. Studies that use supplemental ways of measuring labor market outcomes, such as surveys, are needed to better estimate the effect of correctional education on postrelease employment.

Policy Implications

Our study demonstrates that correctional education improves the chances that inmates who are released from prison will not return and may improve their chances of postrelease employment. Our findings are stable even when we limit our analyses to those studies with more rigorous research designs, and we find a notable effect across all levels of education, from adult basic education and GED programs to postsecondary and vocational education programs. This is important, because the academic needs of inmates are heterogenous. Further, our cost analysis suggests that correctional education programs can be cost-effective. And as noted by Gaes (2008), correctional education is a form of intervention that can affect almost every offender compared with other types of rehabilitative services provided within prisons.

At the same time, it is important to keep in mind that much is changing in the field of correctional education. As noted above, the 2008 recession affected correctional education programs leading to major changes in the number of programs offered, the sizes of classes, the modes of delivery, and the number of inmates who participate in these programs. In addition, the implementation of the new GED exam in 2014 (GED Testing Service, undated)—which will entail a more rigorous test aligned with the Common Core State Standards (CCSS) and computer-based testing (CBT)—will be a new challenge for the field to adjust to and underscores the growing role of computer technology in correctional education.

Going forward, there is a need to undertake studies that get inside the black box to identify the characteristics of effective programs in terms of such elements as curriculum, instructional practices, quality, and dosage. To inform policy and funding decisions at the state and federal levels, policymakers need additional information and a better understanding about how these programs work (and what does not work). In addition to the need for more rigorously designed studies, we also need studies that drill down to examine different aspects of effective programs. For example, understanding how dosage may vary for different types of effective programs would be useful information for administrators and policymakers who are weighing various trade-offs in terms of program duration, frequency, and capacity.

One option is for state and federal policymakers and foundations to invest in well-designed evaluations of correctional education programs to inform such policy questions. Also, researchers and program evaluators need to strive to implement rigorous research designs to examine questions related to potential bias and program dosage. They should ideally strive to measure both proximal and distal outcomes, where the former refers to near-term outcomes, such as test

scores or behavior in prison, and the latter to longer-term outcomes, such as postrelease recidivism and employment. Funding grants and guidelines can help further the field by requiring the use of more rigorous research designs. Such funding also would enable correctional educators to partner with researchers and evaluators to undertake rigorous and comprehensive evaluations of their programs. In addition, a study registry of correctional education evaluations would further aid in the development of the evidence base in this field to help inform policy and programmatic decisionmaking. Knowing that these programs are cost-effective, if these programs were refined based on this important missing information, correctional education might have the potential to yield even greater returns on investment.

APPENDIX A

Document Identification Parameters and Sources

Search Terms

To identify documents for potential inclusion in our analysis, we conducted a search for the phrases "correctional education" and "prisoner education." Additionally, we conducted a search using every potential combination of the following:

1. Academic Term AND Correctional Term
2. Vocational Term AND Correctional Term

Academic Terms

Education
Academic
School
Diploma
GED
Literacy
Math
Reading
Science
College

Vocational Terms

Job skills
Job training
Apprentice
Apprenticeship
Vocational education
Voc-tech
Occupational education
Career and technical education
Workforce development
Workforce training
Workforce preparation
School-to-work

Correctional Terms
Prison
Jail
Incarceration
Inmate
Detention Center
Corrections

Research Data Bases Searched

Education Resources Information Center (ERIC)
Education Abstracts
Criminal Justice Abstracts
National Criminal Justice Reference Service Abstracts
Academic Search Elite
EconLit
Sociological Abstracts
Google Scholar
Rutgers Library of Criminal Justice Grey Literature Database

Online Research Repositories Searched

Vera Institute of Justice
Urban Institute
Washington State Institute for Public Policy
American Institutes for Research
Mathematica Policy Research
John Jay College of Criminal Justice Re-entry Institute
Justice Policy Institute
Center for Law and Social Policy (CLASP)
Juvenile Justice Educational Enhancement Program (JJEEP)
RTI International
Manpower Demonstration Research Corporation (MDRC)

Major Literature Reviews and Meta-Analyses Searched

Steve Aos, Marna Miller, and Elizabeth Drake, *Evidence-Based Adult Corrections Programs: What Works and What Does Not*, Olympia, Wash.: Washington State Institute for Public Policy, #06-01-1201, January 2006.

J. D. Ayers, ed., *The National Conference on Prison Education: Its Role and Practice in the Modern Penitentiary—Proceedings*, Victoria, British Columbia, October 13–15, 1981.

P. E. Barton and R. J. Coley, *Captive Students: Education and Training in America's Prisons*, Princeton, N.J.: Educational Testing Service, Policy Information Center 1996.

J. A. Bouffard, D. L. MacKenzie, and L. J. Hickman, "Effectiveness of Vocational Education and Employment Programs for Adult Offenders: A Methodology-Based Analysis of the Literature," *Journal of Offender Rehabilitation*, Vol. 31, Nos. 1–2, 2000, pp. 1–41.

D. Brazzell, A. Crayton, D. A. Mukamal, A. L. Solomon, and N. Lindahl, *From the Classroom to the Community: Exploring the Role of Education During Incarceration and Reentry*, Washington, D.C.: Urban Institute, 2009.

S. M. Brodus, "Research on Programs in Correctional Institutions," *Justice Professional*, Vol. 14, Nos. 2–3, 2001, pp. 171–200.

C. A. Chappell, "Post-Secondary Correctional Education and Recidivism: A Meta-Analysis of Research Conducted 1990–1999," *Journal of Correctional Education*, Vol. 55, No. 2, 2004, pp. 148–169.

J. P. Conrad and J. Cavros, *Adult Offender Education Programs*, Washington, D.C.: U.S. Department of Justice, National Institute of Justice, Office of Development, Testing, and Dissemination, 1981.

G. G. Gaes, "The Impact of Prison Education Programs on Post Release Outcomes," Reentry Roundtable on Education, March 31 and April 1, 2008.

R. A. Hall and M. W. Bannatyne, "Technology Education and the Convicted Felon: How It Works Behind Prison Walls," *Journal of Correctional Education*, Vol. 51, No. 4, 2000, pp. 320–323.

B. Harrison and R. C. Schehr, "Offenders and Post-Release Jobs: Variables Influencing Success and Failure," *Journal of Offender Rehabilitation*, Vol. 39, No. 3, 2004, pp. 35–68.

M. Jancic, "Does Correctional Education Have an Effect on Recidivism?" *Journal of Correctional Education*, Vol. 49, No. 4, 1998, pp. 152–161.

E. L. Jensen and G. E. Reed, "Adult Correctional Education Programs: An Update on Current Status Based on Recent Studies," *Journal of Offender Rehabilitation*, Vol. 44, No. 1, 2006, pp. 81–98.

S. Lawrence, D. P. Mears, G. Dubin, and J. Travis, *The Practice and Promise of Prison Programming*, Washington, D.C.: Urban Institute, 2002.

R. Linden and L. Perry, "The Effectiveness of Prison Education Programs," *Journal of Offender Counseling, Services, and Rehabilitation*, Vol. 6, No. 4, 1982, pp. 43–57.

D. L. MacKenzie, *What Works in Corrections: Reducing the Criminal Activities of Offenders and Delinquents*, Cambridge, Mass.: Cambridge University Press, 2006, Chapter 5: Academic Education and Life Skill Programs and Chapter 6: Vocational Education and Work Programs.

F. S. Pearson and D. S. Lipton, "Meta-Analytic Review of the Effectiveness of Corrections-Based Treatments for Drug Abuse," *Prison Journal*, Vol. 79, No. 4, 1999, pp. 384–410.

P. Phipps, K. Korinek, S. Aos, and R. Lieb, *Research Findings on Adult Corrections' Programs: A Review*, Olympia, Wash.: Washington State Institute for Public Policy, 1999.

C. Rose, "Women's Participation in Prison Education: What We Know and What We Don't Know," *Journal of Correctional Education*, Vol. 55, No. 1, 2004, pp. 78–100.

R. R. Ross and A. Fabiano Elizabeth, *Time to Think: A Cognitive Model of Delinquency Prevention and Offender Rehabilitation*, Institute of Social Sciences and Arts, 1985.

T. A. Ryan, "Literacy Training and Reintegration of Offenders," *Journal of Correctional Education*, Vol. 46, No. 3, 1991, pp. 1–13.

H. Shrum, "No Longer Theory: Correctional Practices That Work," *Journal of Correctional Education*, Vol. 55, No. 3, 2004, pp. 225–235.

J. M. Taylor, "Should Prisoners Have Access to Collegiate Education? A Policy Issue," *Educational Policy*, Vol. 8, No. 3, 1994, p. 315.

A. Tracy, "Standing Up For Education," *Corrections Today*, Vol. 60, No. 2, 1998, p. 144.

C. B. A. Ubah, "A Critical Examination of Empirical Studies of Offender Rehabilitation-Correctional Education: Lessons for the 21st Century," *Journal of Correctional Education*, Vol. 53, No.1, 2002, pp. 13–19.

J. S. Vacca, "Educated Prisoners Are Less Likely to Return to Prison," *Journal of Correctional Education*, Vol. 55, No. 4, 2004, pp. 297–305.

B. Wade, "Studies of Correctional Education Programs," *Adult Basic Education and Literacy Journal*, Vol. 1, No. 1, 2007, pp. 27–31.

S. A. Ward, "Career and Technical Education in United States Prisons: What Have We Learned?" *Journal of Correctional Education*, Vol. 60, No. 3, 2009, pp. 191–200.

M. Williford, *Higher Education in Prison: A Contradiction in Terms?* Washington, D.C.: National University Continuing Education Association, 1994.

D. B. Wilson, C. A. Gallagher, M. B. Coggeshall, et al., "A Quantitative Review and Description of Corrections-Based Education, Vocation, and Work Programs," *Corrections Management Quarterly*, Vol. 3, No. 4, 1999, pp. 8–18.

D. B. Wilson, C. A. Gallagher, and D. L. MacKenzie, "A Meta-Analysis of Corrections-Based Education, Vocation, and Work Programs for Adult Offenders," *Journal of Research in Crime and Delinquency*, Vol. 37, No. 4, 2000, pp. 347–368.

APPENDIX B

Scientific Review Team Members

Cathryn Chappell, Ph.D., Assistant Professor of Educational Foundations at Ashland University

John Dowdell, M.Ed., Director, Gill Center for Business and Economic Education at Ashland University; Co-Editor of the *Journal of Correctional Education*

Joseph Gagnon, Ph.D., Associate Professor of Special Education at the University of Florida

Paul Hirschfield, Ph.D., Assistant Professor of Sociology and Criminal Justice at Rutgers University

Michael Holosko, Ph.D., Professor of Family and Child Welfare at the University of Georgia

David Houchins, Ph.D., Associate Professor of Educational Psychology and Special Education at Georgia State University

Kristine Jolivette, Ph.D., Associate Professor of Educational Psychology and Special Education at Georgia State University

Larry Nackerud, Ph.D., Professor of Social Work at the University of Georgia

Ed Risler, Ph.D., Professor of Social Work at the University of Georgia

Margaret Shippen, Ph.D., Professor of Special Education and Rehabilitation at Auburn University

APPENDIX C

Meta-Analysis Diagnostic Tests

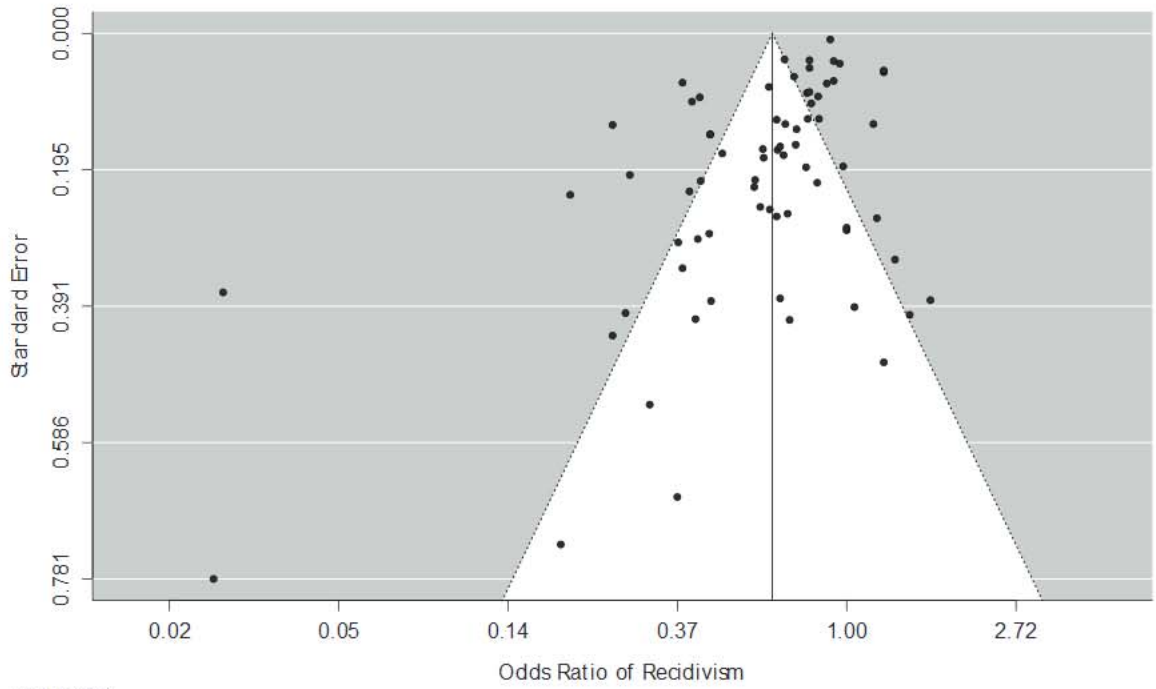
Diagnostic Tests for Recidivism Analysis

If all studies used samples from the same population, we would expect the observed variation in effect sizes to be random, with most (approximately 95 percent) studies' confidence intervals including the pooled effect size of 0.64. The patterning of the boxes and whiskers in Figure 3.1 indicates that this is not the case—and instead suggests that there is substantial heterogeneity in effect sizes that is above the level that would be expected due to random variation. The degree of heterogeneity can be formally assessed through the I^2 statistic, which represents the percentage of variation across studies that is due to heterogeneity, rather than random variation. In this meta-analysis, the value of I^2 is 92 percent, indicating considerable amounts of heterogeneity.

The funnel plot, shown in Figure C.1, is used to look for evidence of publication bias. The funnel plot shows shows each estimate of the odds ratio of recidivating on the x-axis and its standard error on the y-axis. If there were no publication bias, we would expect the points to be approximately symmetrically distributed around the central line, with the spread of points points increasing as the standard error increases. The funnel plot indeed shows some increasing spread with increasing standard errors, but at the larger values of the standard error, the points are no longer distributed symmetrically. This suggests evidence of publication bias, as we would expect smaller studies that found non-significant or negative results to be included in the plot in the lower right half, but these are missing from our search. That this portion of the chart is relatively empty suggests that these studies may exist but have not been published. The Egger regression test of non-symmetricality gives $p < 0.05$. This finding of publication bias suggests that our results may be biased upward (in other words, showing too large an impact on recidivism reduction). However, the publication bias is likely to be small, for three reasons: (1) The number of missing studies is small—the addition of two effect sizes would balance the funnel plot; (2) the missing studies are small and therefore unlikely to have a large effect on our pooled effect size; and (3) two effect sizes are extremely low—such outliers are likely to bias the results of the regression test. The alternative to the Egger regression test is the Begg non-parametric rank test, which is not affected by outliers. In that test, the p-value of 0.450 is non-significant, a finding consistent with no publication bias, though the exact p-value cannot be calculated in the presence of ties.

A second diagnostic test is a leave-one-out analysis. In this analysis, each effect size is sequentially removed from the dataset, and the meta-analysis is rerun. The effect is replaced, and the next effect is removed. This analysis determines the extent to which our results are reliant on one study, and whether our conclusions will be changed with the exclusion of a particular effect. Table C.1 shows the odds ratios and confidence intervals for 70 meta-analyses—with each effect size removed. The table shows that the results are highly stable, and not dependent on any particular study.

Figure C.1
Funnel Plot for Studies of Recidivism



RAND RR266-C.1

Table C.1.
Leave-One-Out Analysis for Studies of Recidivism

First Author (Year)	Odds Ratio	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Adams (1994a)	0.64	0.58	0.69
Adams (1994b)	0.64	0.59	0.69
Adams (1994c)	0.64	0.59	0.69
Allen (2006a)	0.63	0.57	0.69
Allen (2006b)	0.64	0.59	0.69
Anderson (1981)	0.65	0.60	0.70
Anderson (1991)	0.64	0.59	0.70
Anderson (1995)	0.64	0.58	0.69
Batiuk (2005a)	0.64	0.59	0.69
Batiuk (2005b)	0.64	0.59	0.69
Batiuk (2005c)	0.65	0.61	0.71
Batiuk (2005d)	0.64	0.59	0.69
Blackburn (1981)	0.65	0.60	0.70
Blackhawk (1996)	0.64	0.59	0.69

Table C.1. —Continued

First Author (Year)	Odds Ratio	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Brewster (2002a)	0.64	0.59	0.69
Brewster (2002b)	0.63	0.58	0.69
Burke (2001)	0.64	0.59	0.70
Castellano (1996)	0.65	0.60	0.71
Clark (1991)	0.65	0.60	0.70
Coffey (1983)	0.64	0.59	0.69
Cronin (2011)	0.64	0.59	0.70
Davis (1986)	0.63	0.58	0.69
Dickman (1987)	0.64	0.59	0.70
Downes (1989)	0.64	0.59	0.69
Gaither (1980)	0.64	0.59	0.70
Gordon (2003a)	0.66	0.61	0.71
Gordon (2003b)	0.65	0.60	0.70
Harer (1995)	0.64	0.59	0.70
Holloway (1986)	0.64	0.59	0.70
Hopkins (1988)	0.65	0.60	0.70
Hull (2000a)	0.65	0.60	0.70
Hull (2000b)	0.65	0.60	0.70
Johnson (1984)	0.64	0.59	0.69
Kelso (1996a)	0.65	0.60	0.70
Kelso (1996b)	0.65	0.60	0.70
Langenbach (1990)	0.65	0.60	0.70
Lattimore (1988)	0.64	0.59	0.70
Lattimore (1990)	0.64	0.59	0.70
Lichtenberger (2007)	0.64	0.59	0.70
Lichtenberger (2009)	0.64	0.59	0.69
Lichtenberger (2011)	0.64	0.59	0.69
Lockwood (1991)	0.64	0.59	0.70
Markley (1983)	0.64	0.59	0.69
McGee (1997)	0.66	0.61	0.71
Nally (2011)	0.65	0.60	0.70
New York (1992a)	0.64	0.59	0.69

Table C.1. —Continued

First Author (Year)	Odds Ratio	Lower 95% Confidence Interval	Upper 95% Confidence Interval
New York (1992b)	0.65	0.60	0.70
Nuttall (2003)	0.64	0.59	0.69
O'Neil (1990)	0.64	0.59	0.70
Piehl (1995a)	0.64	0.59	0.70
Piehl (1995b)	0.64	0.59	0.70
Ryan (2000)	0.65	0.60	0.70
Saylor (1991)	0.64	0.59	0.70
Schumacker (1990a)	0.64	0.59	0.69
Schumacker (1990b)	0.64	0.59	0.70
Schumacker (1990c)	0.64	0.59	0.70
Smith (2005a)	0.64	0.59	0.69
Smith (2005b)	0.64	0.59	0.69
Smith (2005c)	0.64	0.59	0.69
Smith (2005d)	0.64	0.59	0.69
Steurer (2003a)	0.64	0.59	0.69
Steurer (2003b)	0.64	0.59	0.70
Steurer (2003c)	0.64	0.59	0.70
Torre (2005)	0.65	0.60	0.71
Van Stelle (1995)	0.64	0.59	0.70
Washington (1998)	0.64	0.59	0.69
Werholtz (2003)	0.64	0.59	0.69
Winterfield (2009a)	0.64	0.59	0.70
Winterfield (2009b)	0.65	0.60	0.70
Winterfield (2009c)	0.64	0.59	0.69
Zgoba (2008)	0.64	0.59	0.70

Diagnostic Tests for Employment Analysis

As with the recidivism analysis, the forest plot for the employment analysis (Figure 4.1) shows considerable variation and non-overlapping confidence intervals. The degree of heterogeneity is reflected in the I^2 statistic, which is 90 percent, only slightly lower than for recidivism, again indicating that there is a great deal of heterogeneity between the studies. The funnel plot (Figure C.2) shows that there is a possibility of publication bias, with small studies that have either no effect or a negative effect apparently missing from the dataset. The regression test of non-symmetricality is statistically significant ($p < 0.05$), but the rank test is not ($p = 0.503$). However, as with the recidivism analysis, a small number of studies would balance the graph, and therefore we do not feel that this is likely to indicate substantive bias in our results.

The leave-one-out analysis, presented in Table C.2, shows that the pooled estimate and confidence intervals are not greatly changed by the inclusion or exclusion of any one study.

Figure C.2
Funnel Plot for Studies of Employment

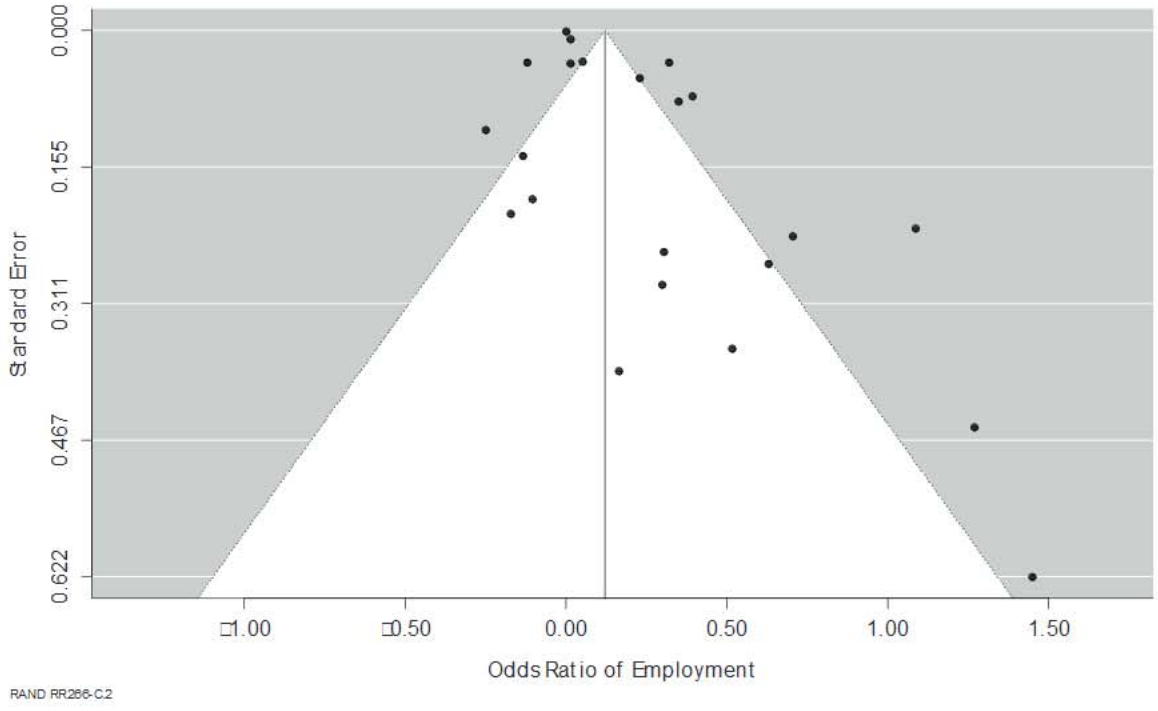


Table C.2
Leave-One-Out Analysis for Studies of Employment

First Author (Year)	Odds Ratio	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Blackhawk (1996)	1.12	1.06	1.19
Cho (2008)	1.19	1.08	1.30
Coffey (1983)	1.12	1.06	1.19
Cronin (2011)	1.09	1.04	1.15
Dickman (1987)	1.14	1.07	1.20
Downes (1989)	1.13	1.07	1.19
Holloway (1986)	1.13	1.07	1.19
Hull (2000)	1.11	1.05	1.17
Lichtenberger (2007)	1.12	1.06	1.18
Lichtenberger (2009)	1.11	1.05	1.18
Sabol (2007a)	1.19	1.09	1.31
Sabol (2007b)	1.16	1.09	1.23
Saylor (1996)	1.11	1.05	1.17
Schumacker (1990a)	1.14	1.07	1.20
Schumacker (1990b)	1.13	1.07	1.19
Schumacker (1990c)	1.12	1.06	1.18
Smith (2005)	1.14	1.08	1.21
Steurer (2003)	1.15	1.08	1.21
Van Stelle (1995)	1.13	1.07	1.20
Visher (2007)	1.13	1.07	1.19
Visher (2011a)	1.15	1.08	1.22
Visher (2011b)	1.14	1.07	1.21

Diagnostic Tests for Computer-Assisted Instruction Analysis

The sample size for the computer-assisted instruction analysis was small, and hence the results of diagnostic tests will be less sensitive. Even pooling across math and reading analyses, the studies were found to be considerably less heterogeneous than studies in the recidivism and employment analyses. I^2 was equal to 0 percent, indicating that there was no greater heterogeneity than would have been expected by chance—and the p-value of the heterogeneity statistic reflected this ($p = 0.435$). The funnel plot in Figure C.3 shows the possibility of some publication bias, with a possible asymmetry in the lower left-hand side; however, the tests of asymmetry were not statistically significant, regardless of whether the regression test ($p = 0.196$) or rank test ($p = 0.180$) was used.

The leave-one-out analysis is presented in Table C.3. It shows that the pooled estimate across math and reading, which is 0.15 grade level equivalents but is not statistically significant (95% CI: –0.05 to 0.35), is not markedly altered by the exclusion of any one study. The confidence interval includes 0, indicating no statistically significant effect, in all cases.

Figure C.3
Funnel Plot for Studies of Computer-Assisted Instruction

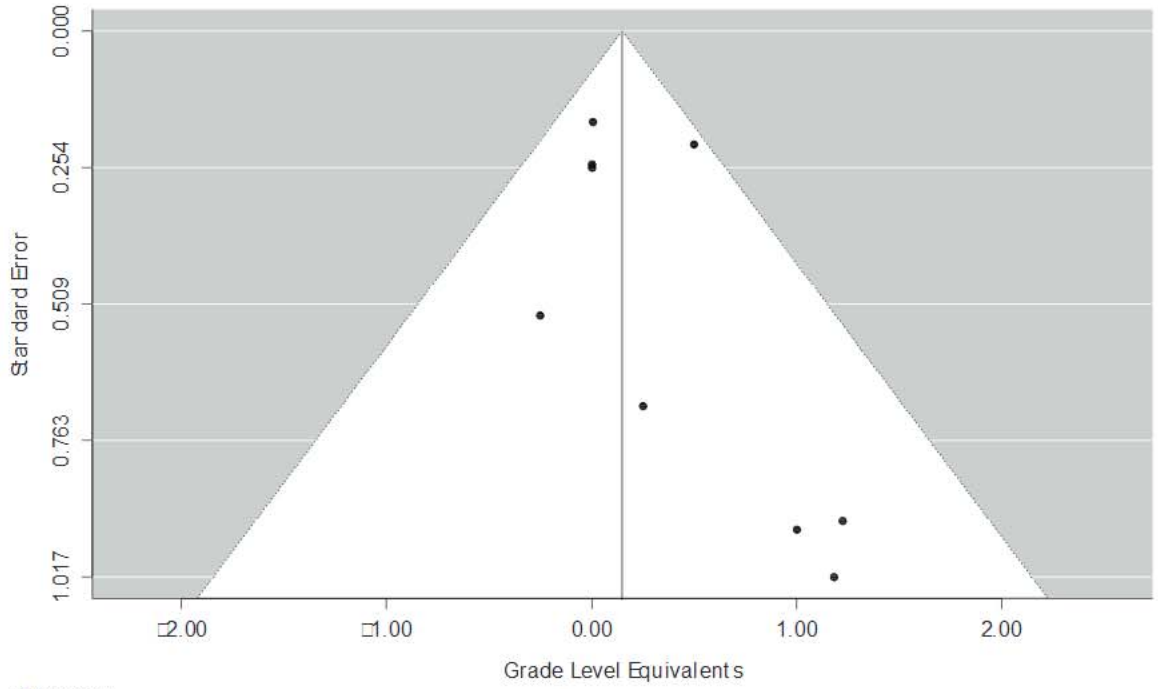


Table C.3
Leave-One-Out Analysis for Studies of Computer-Assisted Instruction

First Author (Year)	Effect Size	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Batchelder (2000a)	0.14	−0.06	0.35
Batchelder (2000b)	0.14	−0.06	0.34
McKane (1996a)	0.14	−0.06	0.34
McKane (1996b)	0.17	−0.04	0.39
McKane (1996c)	0.16	−0.07	0.39
Diem (1980a)	0.23	−0.02	0.48
Diem (1980b)	0.19	−0.05	0.43
Meyer (1983a)	0.19	−0.05	0.43
Meyer (1983b)	0.05	−0.18	0.27

References

Aaronson, D., L. Barrow, and W. Sander, "Teachers and Student Achievement in the Chicago Public High Schools," *Journal of Labor Economics,* Vol. 25, 2007, pp. 95–135.

Angrist, J. D., and J. S. Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion,* Princeton, N.J.: Princeton University Press, 2009.

Aos, S., M. Miller, and E. Drake, "Evidence-Based Adult Corrections Programs: What Works and What Does Not," Washington State Institute for Public Policy, #06-01-1201, January 2006.

Batchelder, J. S., and J. R. Rachal, "Efficacy of a Computer-Assisted Instruction Program in a Prison Setting: An Experimental Study," *Adult Education Quarterly,* Vol. 50, No. 2, February 2000, pp. 120–133.

Baguley, T., "Standardized or Simple Effect Size: What Should Be Reported?" *British Journal of Psychology,* Vol. 100, No. 3, 2009, pp. 603–617.

Bazos, A., and J. Hausman, *Correctional Education as a Crime Control Program,* Los Angeles, Calif.: UCLA School of Public Policy and Research, Department of Policy Studies, March 2004.

Begg, C. B., "Publication Bias," in H. M. Cooper and L. V. Hedges, eds., *The Handbook of Research Synthesis,* New York: Russell Sage Foundation, 1994.

California Rehabilitation Oversight Board, *Biannual Report*, Sacramento, Calif., September 15, 2010.

CASAS, "CASAS Scaled Score References for Grade Levels," n.d. As of December 28, 2012: http://spot.pcc.edu/~shenness/standards/GradeLevelRefs.pdf

Chappell, C. A., "Post-Secondary Correctional Education and Recidivism: A Meta-Analysis of Research Conducted 1990–1999," Electronic Thesis or Dissertation, University of Cincinnati, 2003.

Comprehensive Adult Student Assessment System (CASAS), *Technical Manual,* San Diego, Calif.: 1996.

Cook, T. D., P. M. Steiner, and S. Pohl, "How Bias Reduction Is Affected By Covariate Choice, Unreliability, and Mode of Data Analysis: Results from Two Types of Within-Study Comparisons," *Multivariate Behavioral Research,* Vol. 44, 2009, pp. 828–847.

Crayton, A., and S. R. Neusteter, "The Current State of Correctional Education," Paper prepared for the Reentry Roundtable on Education, New York: John Jay College of Criminal Justice, Prisoner Reentry Institute, 2008.

Cronin, J., "The Path to Successful Reentry: The Relationship Between Correctional Education, Employment and Recidivism," University of Missouri, Institute of Public Policy, September 2011.

Diem, R. A., and P. G. Fairweather, "An Evaluation of a Computer-Assisted Education System in an Untraditional Academic Setting—A County Jail," *AEDS Journal*, Vol. 13 No. 3, Spring 1980, pp. 204–213.

DiIulio, J. J., and A. Piehl, "Does Prison Pay? The Stormy National Debate over the Cost-Effectiveness of Imprisonment," *Brookings Review*, Vol. 9, No. 4, pp. 28–35, 1991.

Egger, M., G. Davey Smith, M. Schneider, and C. Minder, "Bias in Metaanalysis Detected by a Simple, Graphical Test," *British Medical Journal*, Vol. 315, 1997, pp. 629–634.

Farrington, D. P., D. C. Gottfredson, L. W. Sherman, and B. C. Welsh, "The Maryland Scientific Methods Scale," in L. W. Sherman, D. P. Farrington, B. C. Welsh, and D. L. MacKenzie, eds., *Evidence-Based Crime Prevention*, London: Routledge, 2002, pp. 3–21.

Fiedorowicz, C., and R. Trites, *An Evaluation of the Effectiveness of Computer-Assisted Component Reading Subskills Training*, Toronto: Queen's Printer for Ontario, 1987.

Field, A., "Meta-Analysis," in J. N. V. Miles and P. Gilbert, eds., *A Handbook of Research Methods in Clinical and Health Psychology*, Oxford: Oxford University Press, 2005.

Gaes, G. G., "The Impact of Prison Education Programs on Post Release Outcomes," Paper presented at the Reentry Roundtable on Education, John Jay College of Criminal Justice, New York, March 31, 2008.

GED Testing Service, "The New Assessment Is a Stepping-Stone to a Brighter Future," undated. As of June 17, 2013:
http://www.gedtestingservice.com/educators/new-assessment

Gellar, A., I. Garfinkel, and B. Western, "The Effects of Incarceration on Employment and Wages: An Analysis of the Fragile Families Survey," Princeton, N.J.: Princeton University, Center for Research on Child Wellbeing, Working Paper #2006-01-FF, August 2006.

Glaze, L. E., and E. Parks, "Correctional Populations in the United States, 2011," U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, NCJ 239972, November 2012.

Gordon, H. R. D. and B. Weldon, "The Impact of Career and Technical Education Programs on Adult Offenders: Learning Behind Bars," *Journal of Correctional Education*, Vol. 54, No. 4, 2003, pp. 200–209.

Greenberg, E., E. Dunleavy, and M. Kutner, *Literacy Behind Bars: Results from the 2003 National Assessment of Adult Literacy Prison Survey*, Institute for Education Services (IES) National Center for Education Statistics, U.S. Department of Education, NCES 2007-473, May 2007.

Guerino, P., P. Harrison, and W. Sabol, *Prisoners in 2010*, U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, BJS Bulletin NCJ 236096, 2011, revised 2012.

Harer, M. D., *Prison Education Program Participation and Recidivism: A Test of the Normalization Hypothesis*, Washington, D.C.: Federal Bureau of Prisons, Office of Research and Evaluation 1995.

Harlow, C. W., *Education and Correctional Populations*, Bureau of Justice Statistics Special Report, U.S. Department of Justice, NCJ 195670, January 2003; revised April 15, 2003.

Hedges, L. V., E. Tipton, and M. C. Johnson, "Robust Variance Estimation in Meta-Regression with Dependent Effect Size Estimates," *Research Synthesis Methods,* Vol. 1, No. 1, 2010, pp. 39–65.

Hedberg, E. C., "ROBUMETA: Stata Module to Perform Robust Variance Estimation in Meta-Regression with Dependent Effect Size Estimates," Working Paper, 2011.

Henrichon, C., and R. Delaney, *The Price of Prisons: What Incarceration Costs Taxpayers,* Center on Sentencing and Corrections, Vera Institute of Justice, January 2012, updated July 20, 2012.

Hill, C., "Inmate Education Programs," *Corrections Compendium*, Vol. 33, Issue 3, May/June 2008.

Ho, D., K. Imai, G. King, and E. Stuart, "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference," *Political Analysis*, Vol. 15, 2007, pp. 199–236.

Holzer, H. J., S. Raphael, and M. A. Stoll, "Employment Barriers Facing Ex-Offenders," Paper presented at the Urban Institute Re-Entry Roundtable, New York University Law School, May 19–20, 2003.

Kane, T. J., and D. O. Staiger, *Using Imperfect Information to Identify Effective Teachers*, Cambridge, Mass.: National Bureau of Economic Research, April 25, 2005.

Klerman, J. A., and L. A. Karoly, "Young Men and the Transition to Stable Employment," *Monthly Labor Review*, August 1994, pp. 31–51.

Kratochwill, T. R., J. Hitchcock, R. H. Horner, J. R. Levin, S. L. Odom, D. M. Rindskopf, et al., *Single-Case Design Technical Documentation (Version 1.0 [Pilot])*, Washington, D.C.: What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education, 2010.

Kyckelhahn, T., *State Corrections Expenditures, FY 1982–2010*, U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics Bulletin, NCJ 239672, December 2012.

Langan, P. A., and D. J. Levin, *Recidivism of Prisoners Released in 1994*, NCJ 193427, 2002.

Langenbach, M., et al., "Televised Instruction in Oklahoma Prisons: A Study of Recidivism and Disciplinary Actions," *Journal of Correctional Education*, Vol. 41, No. 2, June 1990, pp. 87–94.

Lattimore, P. K., A. D. Witte, and J. R. Baker, *Sandhills Vocational Delivery System Experiment: An Examination of Correctional Program Implementation and Effectiveness*, Washington, D.C.: National Institute of Justice, 1988.

———, "Experimental Assessment of the Effect of Vocational Training on Youthful Property Offenders," *Evaluation Review*, Vol. 14, No. 2, April 1990, pp. 115–133.

Lattimore, P. K., K. Barrick, A. Cowell, D. Dawes, D. Steffey, and S. Tueller, *Prisoner Reentry Services: What Worked for SVORI Evaluation Participants?* Final Report, Prepared for National Institute of Justice, NIJ Grant Number 2009-IJ-CX-0010, Research Triangle Park, N.C.: RTI International and Newark, Del.: University of Delaware, 2012.

Laub, J. H., and R. Sampson, *Shared Beginnings, Divergent Lives: Delinquent Boys to Age 70.* Cambridge: Harvard University Press, 2003.

LeLorier, J., G. Gregoire, A. Benhaddad, J. Lapierre, and F. Derderian, "Discrepancies Between Meta-Analyses and Subsequent Large Randomized, Controlled Trials," *New England Journal of Medicine*, Vol. 337, No. 8, 1997, pp. 536–542.

Lichtenberger, E. J., *The Impact of Vocational Programs on Post-Release Outcomes for Vocational Completers from the Fiscal Year 1999, 2000, 2001 and 2002 Release Cohorts*, Richmond, Va.: Center for Assessment, Evaluation, and Educational Programming, Virginia Polytechnic Institute and State University, 2007.

Levitt, S. D.,"The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Legislation," *Quarterly Journal of Economics*, Vol. 111, 1996, pp. 319–351.

Lipsey, M. W., "The Primary Factors That Characterize Effective Interventions with Juvenile Offenders: A Meta-Analytic Overview," *Victims and Offenders*, Vol. 4, 2009, pp. 124–147.

Lipton, D. S., R. Martinson, and J. Wilks, *The Effectiveness of Correctional Treatment: A Survey of Treatment Valuation Studies*, New York: Praeger Press, 1975.

MacKenzie, D., *What Works in Corrections: Reducing the Criminal Activities of Offenders and Delinquents,* New York: Cambridge University Press, 2006.

———, *Structure and Components of Successful Educational Programs,* Reentry Roundtable on Education, New York, March 31 and April 1, 2008.

Martinson, R. "What Works?—Questions and Answers About Prison Reform," *The Public Interest*, Spring, 1974, pp. 22–54.

McCaffrey, D. F., G. Ridgeway, and A. R. Morral, "Propensity Score Estimation with Boosted Regression for Evaluating Causal Effects in Observational Studies," *Psychological Methods*, Vol. 9, No. 4, 2004, pp. 403–425.

McGlone, J., *Status of Mandatory Education in State Correctional Institutions,* Washington, D.C.: U.S. Department of Education, 2002.

McKane, P. F., and B. A. Greene, "The Use of Theory-Based Computer-Assisted Instruction in Correctional Centers to Enhance the Reading Skills of Reading-Disadvantaged Adults," *Journal of Educational Computing Research,* Vol. 15, No. 4, 1996, pp. 331–344.

Meyer, L. A., J. C. Ory, and R. C. Hinckley, *Evaluation Research in Basic Skills with Incarcerated Adults,* Champaign, Ill.: University of Illinois Center for the Study of Reading, Technical Report No. 303, 1983.

Murnane, R. J., and J. B. Willett, *Methods Matter: Improving Causal Inference in Educational and Social Science Research,* New York: Oxford University Press, 2011.

Myers, D. E., and M. Dynarski, *Random Assignment in Program Evaluation and Intervention Research: Questions and Answers,* NCEE 2003-5001. Washington, D.C.: National Center for Education Evaluation, Institute of Education Sciences, U.S. Department of Education, 2003.

Nally, J., S. Lockwood, T. Ho, and K. Knutson, "The Effect of Correctional Education on Postrelease Employment and Recidivism: A 5-Year Follow-Up Study in the State of Indiana," working paper, 2011.

Pager, D., "The Mark of a Criminal Record," *American Journal of Sociology,* Vol. 108, No. 5, 2003, pp. 937–975.

Pastore, A. L., and K. Maguire, *Sourcebook of Criminal Justice Statistics, 2001*, U.S. Department of Justice, Bureau of Justice Statistics, NCJ 196438, 2002.

Petersilia, J., *When Prisoners Come Home: Parole and Prisoner Reentry*, New York: Oxford University Press, 2003.

———, *Understanding California Corrections*, Berkeley, Calif.: California Policy Research Center, University of California, May 2006.

Pew Center on the States, *State of Recidivism: The Revolving Door of America's Prisons,* Washington, D.C.: The Pew Charitable Trusts, April 2011.

Piehl, A. M., "Learning While Doing Time," Kennedy School Working Paper #R94-25, 1995.

Public Law 110-199, Second Chance Act of 2007, April 9, 2008.

Raphael, S., "The Employment Prospects of Ex-Offenders," *Focus*, Vol. 25, No. 2, Fall–Winter 2007–08.

Rivkin, S. G., E. A. Hanushek, and J. F. Kain, "Teachers, Schools, and Academic Achievement," *Econometrica,* Vol. 73, No. 2, March, 2005, pp. 417–458.

Rockoff, J. E., "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data," *American Economic Review,* Vol. 94, No. 2, 2004, pp. 247–252.

Rosenbaum, P. R., and D. B. Rubin, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, Vol. 70, No. 1, 1983, pp. 41–55.

Rubin, D. B., "Estimating Causal Effects from Large Data Sets Using Propensity Scores," *Annals of Internal Medicine*, Vol. 127, No. 8 (part 2), 1997, pp. 757–763.

Sabol, W. J., "Local Labor-Market Conditions and Post-Prison Employment Experiences of offenders Released from Ohio State Prisons," in Shawn Bushway, Michael A. Stoll, and David F. Weiman, eds., *Barriers to Reentry? The Labor Market for Released Prisoners in Post-Industrial America,* New York: Russell Sage Foundation, 2007, pp. 257–303.

Saylor, W. G., and G. G. Gaes. "Prep: Training Inmates through Industrial Work Participation, and Vocational and Apprenticeship," *Corrections Management Quarterly,* Vol. 1, No. 2, 1996.

Schochet, P. Z., T. D. Cook, J. Deke, G. Imbens, J. R. Lockwood, J. Porter, et al., *Standards for Regression Discontinuity Designs (Version 1.0 [Pilot]),* Washington, D.C.: What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education, 2010.

Seftor, N. S., and D. P. Mayer, "The Effect of Alternative Certification on Student Achievement: A Literature Review: Final Report," Mathematica Policy Research, Inc., 2003.

Shadish, W. R., T. D. Cook, and D. T. Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Boston: Houghton Mifflin, 2002.

Sherman, L. W., D. C. Gottfredeson, D. L. MacKenzie, J. Eck, P. Reuter, and S. Bushway, *Preventing Crime: What Works, What Doesn't, What's Promising*, Washington, D.C.: U.S. Office of Justice Programs, 1997.

Slavin, R. E.,"Meta-Analysis in Education: How Has It Been Used?" *Educational Researcher*, Vol. 13, No. 8, 1984, pp. 6–15.

Stephan, J. J., *Census of State and Federal Correctional Facilities, 2005*, U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, National Prisoner Statistics Program, NCJ 222182, October 2008.

Steurer, S. J., L. G. Smith, and A. Tracy, *Education Reduces Crime: Three-State Recidivism Study*, Lanham, Md.: Correctional Education Association, 2003.

Tolbert, M., *A Reentry Education Model: Supporting Education and Career Advancement for Low-Skill Individuals in Corrections,* MPR Associates, Inc., Prepared for the U.S. Department of Education Office of Vocational and Adult Education (OVAE), 2012.

Torre, M. E., and M. Fine, "Bar None: Extending Affirmative Action to Higher Education in Prison," *Journal of Social Issues*, Vol. 61, No. 3, 2005, pp. 569–594.

Travis, J., A. L. Solomon, and M. Waul, *From Prison to Home: The Dimensions and Consequences of Prisoner Reentry,* Washington, D.C.: Urban Institute Press, 2001.

Uggen, C., "Work as a Turning Point in the Life Course of Criminals: A Duration Model of Age, Employment, and Recidivism," *American Sociological Review*, Vol. 65, 2000, pp. 529–546.

U.S. Census Bureau, Annual Survey of State Government Finances. As of July 26, 2013: http://www.census.gov/govs

U.S. Department of Education, *Transforming American Education: Learning Powered by Technology: National Education Technology Plan 2010*, Washington, D.C.: Office of Educational Technology, U.S. Department of Education, 2010.

Visher, C. A., and P. K. Lattimore, "Major Study Examines Prisoners and Their Reentry Needs," *NIJ Journal*, Vol. 258, 2007, pp. 30–33.

Wells, R. E., "Education as Prison Reform: A Meta-Analysis," unpublished dissertation, Baton Route, La.: Louisiana State University, 2000.

Wertz, J., *Cuts in Inmate Education May Cost Oklahoma Taxpayers Later,* StateImpact Oklahoma, June 27, 2012.

Western, B., J. R. Kling, and D. F. Weiman, *The Labor Market Consequences of Incarceration,* Princeton, N.J.: Princeton University Industrial Relations Section, Working Paper No. 450, January 2001.

What Works Clearinghouse, *Procedures and Standards Handbook (Version 2.1),* Washington, D.C.: Institute of Education Sciences, U.S. Department of Education, 2011.

Wilson, D. B., C. A. Gallagher, and D. L. Mackenzie, "A Meta-Analysis of Corrections-Based Education, Vocation, and Work Programs for Adult Offenders," *Journal of Research in Crime & Delinquency*, Vol. 37, No. 4, 2000, pp. 347–368.

Windham School District, *2011–2012 Budget and Salary Schedule*, Texas Department of Criminal Justice.

Winterfield, L., M. Coggeshall, M. Burke-Storer, V. Correa, and S. Tidd, *The Effects of Postsecondary Correctional Education: Final Report,* Washington, D.C.: Urban Institute, May 2009.

ISBN 978-0-8330-8108-7